DOCUMENT RESUME

ED 314 035 IR 014 122

AUTHOR	Jernstedt, G. Christian; And Others
TITLE	Development and Evaluation of an Educational Data Acquisition System for Use in Secondary Schools. Technical Note 411.
INSTITUTION	Creare, Inc., Hanover, NH.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Sep 86
CONTRACT	400-84-0006
NOTE	78p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Computer Assisted Instruction; *Databases; *Information Retrieval; Intermode Differences; Microcomputers; Pilot Projects; *Science Instruction; *Science Laboratories; Secondary Education

ABSTRACT

The purpose of this research was to design, create a prototype, and evaluate an Educational Data Acquisition System (EDAS) for use in secondary school science laboratories. The prototype of EDAS consisted of a microcomputer, data acquisition interface, sensors, software, and documentation. It provided the teacher and students with a general-purpose linkage between physical processes and the microcomputer and is comparable to the tools routinely employed in industry. The prototype system was tested for an entire school year in a local secondary school. An overview of the technical design and benefits of EDAS is provided along with a description of how the evaluation was performed. (EDAS was tested in a total of 19 different experiments where some class sections were taught using EDAS and other sections were taught using traditional methods.) The evaluation results confirm that EDAS is a powerful tool that can improve learning in crucial ways. (28 references) (Author/GL)

TECHNICAL NOTE 411

...OPMENT AND EVALUATION OF AN ...DATA ACQUISITION SYSTEM ...USE IN SECONDARY SCHOOLS

...ber 1986 TN-411

...ork upon which this publication is
... was performed pursuant to Contract
...84-0006 of National Institute of
...tion. It does not, however, nec-
...lly reflect the view of that agency.

U.S. DEPARTMENT OF EDUCATION

SMALL BUSINESS INNOVATION RESEARCH PROGRAM
PHASE II—FY 84
PROJECT SUMMARY

Topic No. 2

Name and Address of Proposer

Creare, Inc.
P.O. Box 71
Hanover, NH 03755

Name and Title of Principal Investigator

Paul H. Rothe, Senior Engineer and Vice President

Proposer's Title

Senior Engineer and Vice President

Technical Abstract (Unclassified) (Limit To Two Hundred Words)

The purpose of our Phase II research was to design, prototype and evaluate an Educational Data Acquisition System (EDAS) for use in secondary school science laboratories. The prototype EDAS consisted of a micro-computer, data acquisition interface, sensors, software and documentation. It provided the teacher and students with a general-purpose linkeage between physical processes and the microcomputer and is comparable to the tools routinely employed in industry. The prototype system was tested for an entire school year in a local secondary school.

This report provides an overview of the technical design and benefits of the EDAS, and a description of how the evaluation was performed. The EDAS was tested in a total of 19 different experiments where some class sections were taught using the EDAS and other sections were taught using traditional methods. The evaluation results confirm that the EDAS was a powerful tool that can improve learning in crucial ways.

Anticipated Benefits/Potential Commercial Applications of the Research or Development

Commercially, the EDAS would be a turnkey system of components and software immediately useful in all science laboratories of secondary schools, vocational schools, and community colleges. We believe that the introduction of an EDAS will improve laboratory instruction, technical skill development, and the teaching of science, which in turn is essential for our nation to compete in an increasingly technical world.

Key Words

Computers, Science, Education, Laboratory, Computer Aided Instruction (CAI), Computers in Science, Educational Data Acquisition, Laboratory Interfacing

Creare

DEVELOPMENT AND EVALUATION OF AN EDUCATIONAL DATA ACQUISITION SYSTEM FOR USE IN SECONDARY SCHOOLS

Principal Investigator

Paul H. Rothe

Authors

G. Christian Jernstedt

Tyrone D. Cannon

Paul H. Rothe

Susan A. Schwarz

Creare, Inc.
Hanover, NH

TN-411
1 September 1986

Creare

EXECUTIVE SUMMARY

Overview

This report describes the results of a project sponsored by the Department of Education Small Business Innovation Research Program. In March, 1984, Creare completed the first phase of this project, the results of which were presented in a final report entitled, "Feasibility Study of An Educational Data Acquisition System for Use in Secondary Schools."

We began Phase Two in July, 1985, under contract number 400-84-0006. We designed, prototyped, and then evaluated an Educational Data Acquisition System (EDAS) in a local high school. The objectives of the Phase Two project were:

1. define the range of educational needs for EDAS,

2. establish technical descriptions of the educational data acquisition system, and

3. evaluate student and teacher response to automated laboratory science.

All of these objectives have been achieved in the Phase Two project as discussed, in turn, below.

Activities

In Phase One and at the beginning of the Phase Two project we reviewed the existing data acquisition products in use in schools and the acceptance of these devices by the schools. Their usage was very low and by examination of the products themselves, talking to teachers who were and were not familiar with data acquisition and surveying the literature on computer-aided instruction in the science classroom, we developed a definition of schools' requirements. These requirements were reviewed by our educational consultants and were then used in our design of the EDAS hardware and software and in our evaluation criteria for the EDAS.

During the first four months of Phase Two we performed the technical design of the EDAS hardware and software components. The design incorporated the concepts that Creare uses for its industrial data acquisition systems, results from our market research on microcomputer-based data acquisition systems, and the educational requirements from our surveys. We spent the next six months developing the prototype system, followed by a trial period of two months of preliminary usage and evaluation in the school. We then spent three additional months updating the prototype to incorporate improvements suggested by students, teachers, and evaluator feedback. The prototype was then placed in the school for a year of testing. The effectiveness of the design implemented in the prototype EDAS led to a successful

Creare

evaluation and favorable response of students and teachers across a wide variety of subject areas and grade levels.

The EDAS evaluation lasted for an entire academic year. A research assistant in permanent residence at the high school obtained classroom observational data, information about student and teacher attitudes, and the results of students' written work for all subject areas and grade levels throughout the year. The impact of the EDAS and the benefits and pitfalls of automated data collection in the science classroom learned from the evaluation will have major significance in future development of data acquisition products for the classroom.

Organization of the Report

The following three chapters present the results of the Phase Two project. These chapters are organized as a series of three professional articles about the EDAS. Each chapter has been or is being submitted to refereed journals for publication. While this plan of publication ensures wide communication of our results it does result in some unavoidable, but minimal, redundancies across the three chapters for the reader of this report. Chapter I presents an overview of both the technical and evalu ive portions of the project, emphasizing the benefits of the EDAS. The analysis of the data in Chapter I is based on two case studies from the 19 experiments that were run. This type of analysis can sometimes illuminate information that is not revealed by the overall analyses of Chapter III, but can miss information that the methods of Chapter III can detect. Chapter II describes the details about how the educational evaluation was set up and operated. It is designed to serve both as an archival record of what we did and as a manual for others who wish to apply similar evaluation methods to other problem areas. Chapter III presents the detailed results of the evaluation. We employed innovative methods in Chapter III to examine simultaneously the results of 19 different experiments.

Conclusion

We have met all the requirements of the project. We developed a prototype EDAS for use in secondary science classes and had it rigorously evaluated by educational researchers. The results show that computerized data collection in the classroom will foster excellence in the sciences and allow students to achieve their full potential.

Creare

ACKNOWLEDGEMENTS

The EDAS was developed by a team of individuals. Paul H. Rothe, of Creare, Inc., conceived of the project and was the project director. Susan A. Schwarz, at Creare, Inc., was the project engineer and directed the technical development. G. Christian Jernstedt, at Dartmouth College, directed the educational development and research. Tyrone D. Cannon, at Dartmouth, was the research assistant and field manager for the educational research.

The authors wish to acknowledge the individual contributions of the team of professionals and the many others who contributed to the thinking that forms the basis of this report. So many individuals gave so much of their time to this project. Its quality reflects their contributions.

John Hawkins, Karen Sessler, and Matthew Miller of Creare skillfully carried out the software development with a keen insight into the needs of the users and a desire and ability to create a truly exceptional software package.

Robert Eccher, at Creare, was responsible for sensors and signal conditioning. His experience, creativity, and effort is what helped make the EDAS aware of the world which the students were trying to understand.

The science department at Hanover High School was instrumental in the success of the project. Rob Buckley, Warren DeMont, Barbara Hirai, John Hutchins, Bruce Koloseike, Dick Murphy, Dale Rowe, and Peter Weimersheime each contributed both a vision of excellence and the hard work necessary to achieve it. They are exceptional teachers.

The nearly 500 students enrolled in science courses at Hanover High School met our many requests for data from them with occasional groans but an everpresent willingness to help and a deep curiosity about their school and the learning process.

Uwe Bagnato, Principal of Hanover High School, had the confidence in his faculty and students and the understanding of the importance of feedback in effective teaching and learning to open the school to our evaluation efforts.

Eugenia F. Braasch, at Dartmouth, brought enormous energy to carrying out the special analyses in the evaluation and provided critical insights into the interpretation and presentation of the results so that others might understand and use them.

Amalia Klinger, at Dartmouth, conducted exhaustive and sometimes exhausting literature reviews with a keen eye to the nuances in the work of the many investigators of computer based instruction and classroom learning. She gave many hours of great care to the entry and validation of the huge data base about the students who participated in the study.

Creare

TABLE OF CONTENTS

Creare

CHAPTER I: OVERVIEW OF THE PROJECT

A Collaborative Project to Develop Computer Based Instruction in Science Laboratory Courses

by G. Christian Jernstedt
and
Paul H. Rothe

This chapter describes the results of an intensive three year program to bring the sophistication of industrial technology to the aid of science teaching and learning. With support from the Department of Education a team of experts was assembled from Creare, Inc., an advanced engineering consulting company specializing in computer software and systems, from the Psychology Department at Dartmouth College, a department with specialists in learning, curriculum design, and evaluation, and from Hanover High School, an excellent regional school with a six person science department that offers the basic scientific disciplines for a range of student ability levels.

The team had two goals. The first goal was to take the best ideas for improving science education with computer based tools and design and create an Educational Data Acquisition System (EDAS). The second goal was to evaluate the effectiveness of the EDAS in actual class use in a school system, asking (a) will the system work as designed, (b) can and will teachers use it in class, and (c) will the system lead to desirable changes in students' attitudes and performance?

The Design Criteria for the Operation of the EDAS

In laboratory science computers are a critical element of modern research. Computers play an essential role for laboratory scientists, engineers, physicians, office workers, executives, factory workers, artists; the list is nearly endless. The core of the EDAS was designed around a personal computer to address the broad need for true computer literacy in science while improving the educational process as well. Although it was primarily designed for and was tested in secondary school classrooms, the EDAS can theoretically be used with students of nearly any age.

Creare

The overall goal in developing the specific apparatus was to provide flexible tools for the teachers to use in developing curriculum rather than rigidly fixed or required laboratory instruments and procedures. At Creare specialists in software, hardware, and documentation designed the computer system. To assist in answering questions about what works in the classroom and what teachers will accept and use, the design process began with an information review that identified the existing computer data acquisition devices in use in the schools and their acceptance by teachers. A survey in 1982 of 1200 schools showed that only 12% were using computers in science classes (Harvard, 1982). Only 2% of that 12% (e.g., 3 schools, or 0.2% overall) actually used a laboratory interface in their curriculum. Reviews of such software and hardware (Ahl, 1983), product literature, conversations with teachers who use the products, and hands-on evaluation of some of the products indicated that there were good reasons why use of such interfaces was low. The problems uncovered included the relatively passive role of the students in using some of the instruments, the fixed, limited nature of what can be done with the instruments, the absence of productive software to employ effectively the instruments, and the very limited, almost game-like quality of much of the apparatus.

Based on the research into computers in science, the EDAS design criteria differed from those of most current school instrumentation products in that the criteria specified a general purpose, industrial quality, data acquisition and reduction system. The school environment is probably harder on instruments than is the research laboratory. We decided that students and teachers need the same quality and ease of use that is available in industry; they need equipment and curriculum that emulates industrial scientific research. The EDAS should unobtrusively guide students and teachers into effective scientific activities. Toward that end, a set of eight features was designed for the EDAS:

- fast,
- accurate,
- patient,
  dedicated,
- efficient,
- general purpose,
  modular,
- comprehensive,
  uniform in operation,
- orderly,
  flexible,
- simple to use, and
  easy to use.

The first four features of the EDAS differentiate it from equipment used by students in a manual, traditional experiment. The speed of the EDAS should allow students to exploit sensors and investigate processes that they could never consider in the past, while its patience

Creare

allows it to perform large numbers of repetitive operations with quick, efficient instructions given by the student. By being dedicated to a measurement task the EDAS can collect data from long-term experiments, such as weather sampling over the period of days, weeks, or longer.

The single most important feature of the EDAS is that it is general purpose in its application. Teachers and students vary widely in their abilities and familiarities with computing and science. The teacher and student should not be constrained to pre-programmed or canned laboratory work. They must have the power of a tool that serves their own, individual needs, and they should be encouraged to explore and define those needs. The modular nature of the EDAS allows it to be configured easily for different science courses. In a moment a new sensor module can be attached, changing the EDAS from a biological to a chemical instrument. Yet, the software interface and general operating procedures remain constant, serving to unify experiments in various disciplines. We also envisioned moving the EDAS from one classroom to another during the class change period, which would allow it to be used more or less continuously and thus significantly reduce its cost per classroom.

The EDAS should be comprehensive in that it should contain all of the elements necessary to conduct an experiment, from data acquisition to final evaluation. The user must be freed to think about the data and their meaning, not the logistics of idiosyncratic, often obsolete apparatus. Employing an orderly, but open menu approach guides the user in developing an orderly process of experimentation. Since the menu provides options, not requirements, teachers may tailor activities to the special needs, abilities, and interests of students. Analytically minded students might spend a greater proportion of their time with data handling functions, while mechanically minded students might spend more time in the actual conducting of the experiment.

The ease of use should make exploration with the EDAS simple, even fun. We wanted a system that would disappear into the background as the student thought creatively and openly about the process that he or she was investigating. We hoped for lots of "what if?" questions from the users, who would then proceed to answer the questions experimentally. As the system is shared across courses, students should build a common understanding of scientific method and be encouraged to bring what they have learned in other science courses into each new one.

The Design Criteria for the Educational Impact of the EDAS

We identified five critical features of effective educational environments that must be present in the science classroom and that must be supported by the use of the EDAS. As they form the educational foundation on which the EDAS development rested, they will be reviewed in some depth.

Creare

Interactive

an active learner

a responsive environment

an orientation towards retrieval processes

creation of a base of experience on which to build later learning

Accessible

attention getting

challenging

flexible

cognitively structured

efficient

Collaborative

Personalized

use of familiar curricula

manageable unit format

flexible pacing

mastery criteria

Generalizable

The EDAS is an **interactive** tool. Students who are more active in their engagement with their subject matter and their classroom peers learn more (Webb, 1983). Furthermore we know that the study activities of such students serve as a catalyst to increase the effects of their motivation and ability (Jernstedt & Chow, 1980). Due to its speed, efficiency, and patience, the computer in the EDAS is always waiting for student action. This implicit demand, characteristic of the computer, is not seen by students as pressure so much as positive incentive to engage, as evidenced by the computer's positive effect on classroom attitudes (Kulik, Bangert, & Williams, 1983).

Generally, as student work is monitored and feedback is provided, learning improves (Berliner & Rosenshine, 1977). When students work in groups, however, teachers cannot provide effective individual monitoring of student activities. The ability to monitor

performance and respond immediately, directly, and in a manner that is focused on the important questions is a necessity in science courses where long sequences of steps may be taken before an answer, the feedback to the student, is achieved. The EDAS will thus serve a responsive role in the classroom by directly requesting and reacting to students' actions, such as their connecting of sensors and entering of information, data, or commands.

One of the most basic needs in the classroom is for an orientation on the retrieval (rather than the storage) mental processes of the learners. Learners must rehearse what they have learned; it is the crucial learning process, and we have known about it for many years (Spitzer, 1939). As described earlier, it is not what learners see or hear, but what they do that most determines what they learn. This retrieval focus is aided by learning materials and activities that are structured, provide imagery, and encourage elaboration.

Let us consider structure first. The science curriculum can be described in terms of lower and higher order rules. The use of a few well chosen higher order rules can drastically reduce the number of lower order rules that a student must learn. Moreover, the students will perform just as well on problems that require the use of the lower order rules that were eliminated (Scandura, 1977). The modular and general purpose design of the EDAS is based not only on economic foundations, but also on the desire to present the concepts of laboratory science as a relatively few, broadly applicable principles of data collection, and analysis, and presentation. The fact that students will encounter a common set of high level tools in each laboratory exercise ensures mastery of those major principles.

The retrieval process is also aided by the use of imagery. Much of current educational software, even for high school science, has its roots more in the imagery of comic books than of scientific illustration. Scientific graphic image presentation, of the form present in the EDAS, can facilitate learning and improve retention (Wittrock & Lumsdaine, 1977).

The asking of questions that require deep intellectual processing of information can be done almost ideally in the science laboratory (Craik & Lockhart, 1972). Traditional science exercises are often heavy on manual work while light on intellectual activity. The relative efficiency with which hypotheses can be tested with the EDAS can minimize the drudging routine activities and hence encourage a greater degree of thought and questioning, both on the part of teachers of their students and by students themselves. The result of such activities is increased retention (Horton & Mills, 1984).

More than any other area of the curriculum, science requires direct student experience. There is a vast gulf between the actions of describing a scientific principle and those of correctly using that principle (Simon, 1980). If students are to think like scientists they must begin to learn how to act like scientists. The experienced individual views events and ideas in ways that are fundamentally different from those of the inexperienced individual (Reed & Johnson, 1977). As students with the EDAS increase the proportion of their time that is spent in intellectual work rather than clerical activities, they will acquire more extensive experience in the fundamentals of creative science.

Though they may be well designed, learning materials are of minimal effectiveness if they are not perceived as accessible to student and to teacher. This accessibility can be improved through using science apparatus that is attention getting, challenging, flexible, cognitively structured, and efficient.

Creare

Computers in the classroom can produce a higher level of student attention (Jernstedt, 1983). Furthermore, increased attentiveness is significantly correlated with achievement. Because the EDAS provides a familiar laboratory interface for a wide variety of experiences, it will avoid the "drowning out" effect that accompanies the use of continually novel materials (Piontkowski & Calfee, 1979).

Though science is a lively intellectual area, with changes in knowledge occurring daily, science laboratories often deal with topics that have been resolved for decades, if not for centuries. The excitement of open experimental exploration with the EDAS should challenge students to search for information and induce more rehearsal of what is happening in the laboratory (Smith, Johnson, & Johnson, 1981). The flexible, open-ended nature of the "what if" questions that are possible with the EDAS, compared to the specific question answering that is typical of traditional science laboratory experiments, should support long term retention of the material.

Cognitive structure is an additional design criterion that must be considered carefully, since increased, appropriate cognitive effort increases learning and retention (Horton & Mills, 1984). Because the user must select each step of the experimental process from EDAS menus, the EDAS cannot be operated without knowing what one is trying to do; thus the EDAS menus are designed to require intellectual effort.

We have discussed accessibility to the student, and turn now to accessibility to the teacher. The activities of a teacher in a subject matter such as science are quite complex and resemble the skills needed to run a small business (Duke, 1979). But rarely are teachers supported with the tools that are typically available in such a business. The efficiency of the computer in handling many different students and many different experiments while producing conceptually uniform reports will help teachers achieve some of the productivity gains that computers are bringing to the business world. The increased teacher productivity can produce other desired effects. In place of the time normally spent for laboratory set up, take down, and other such manual work the teacher's job will be to frame questions whose answers will illustrate the critical principles of the subject matter. The teacher can focus on the intellectual matters more than the logistical ones.

The superior learning that occurs with **collaborative** efforts among students has been known for some time and studied by many investigators (Johnson & Johnson, 1975; Johnson, Maruyama, Johnson, Nelson, & Skon, 1981; Sharan, 1980). Within a class or student lab group the various tasks involved in using the EDAS can be apportioned according to student skill level and interests. One of the strengths of collaborative learning is that each member of a group can feel that he or she is a valued contributor to the group process.

Science education, to be effective, must be **personalized** for both the individual teacher and the individual student. A major problem with federally supported curriculum materials for science education in the past has been the resistance of teachers to fixed, new curricula (Welch, 1979). This problem can be ameliorated with the EDAS by allowing the teacher to combine his or her own mix of self-chosen, familiar materials with a set of standard EDAS operating procedures.

Creare

For the student, personalization focuses largely on dividing material into manageable units that can be mastered at a rate appropriate to the level and background of the student (Ryan, 1974). The goal of small units that can be completed within one class period can be achieved with the speed of the EDAS (Semb, 1974). When units are small, the pacing of a course becomes less troublesome and more flexible. Some students may complete two units while others complete only one.

The use of a mastery criterion is the most common characteristic of personalized courses of instruction. Requiring students to know fundamental skills well, before moving on the more advanced skills, is of critical importance in science education. The research is clear that mastery produces better outcomes (Hursh, 1976). With the EDAS students can easily and quickly repeat experiments, with new parameters, until they master the underlying concepts.

Effective retention of what has been learned in the science laboratory depends in part on the degree to which cues in the new environment generalize because they are similar to those in the original learning environment (moscovitch & Craik, 1976). For this reason, if students are to take what they learn in the high school science course with them in later academic and work environments, those environments must be similar. Computers are an essential part of modern industrial laboratories, and, very recently, of university science laboratories. The degree to which students encounter them in realistic ways in high school will determine the degree to which their high school course is generalizable. In other areas it is clear, and may be so in the sciences as well, that increased exposure to instructional computers can improve attitudes toward the use of computers as tools and toward the courses in which the computers are employed (Kulik, et al, 1983).

The Basic EDAS

The structure of the EDAS is illustrated in Figure I.1. The host microcomputer is a standard IBM PC with monitor, disk drive, serial interface, and printer. The machine has far more power than other classroom computers and, with the advent of IBM clones, is quite affordable. The electronic interface (DAQ) is a Burr-Brown 3002 Data Acquisition System containing an analog to digital converter and its own microprocessor and memory. Signal processors within the DAQ condition the data received from various sensors. The sensors allow the DAQ to sample essentially any physical process, including temperature, pressure, force, motion, pH, conductivity, sound, light, and electricity.

The key to the effectiveness of the EDAS for the user is the software that operates the system. Much of the project effort focused on developing software that was truely flexible and easy to use. Menus offer the user a full range of options: (a) prepare for an experiment, including changing the experimental setup by adding or deleting sensors and checking the status of sensors; (b) run an experiment, including setting the operating mode to storage of data or monitoring and display of data, setting the sampling rate, setting the display format to graphical or tabular, and collecting data; (c) reduce and display data, including selecting data files and displaying their contents; (d) manage the data disks, including cataloging, deleting, and copying files; (e) change the configuration of the system, including specifying the computer and peripherals present and the types of sensors; and (f) receive help. No knowledge of programming or computer operation is required other than the ability to turn the machine on. The student scientist at the control of the machine thinks about the experiment, not the

Creare

computer. The software and system operation are further designed so that teams of students can use the machine, with each performing a different part of the experimental procedure.

A full explanatory manual was written for the EDAS as part of its development. We forsee the eventual creation of other supporting materials including a teacher's guide to the pedagogical aspects of the EDAS, individual guides for a wide variety of laboratory experiments, and a newsletter for the sharing of other ideas about the effective use of the EDAS in the classroom.

Creare

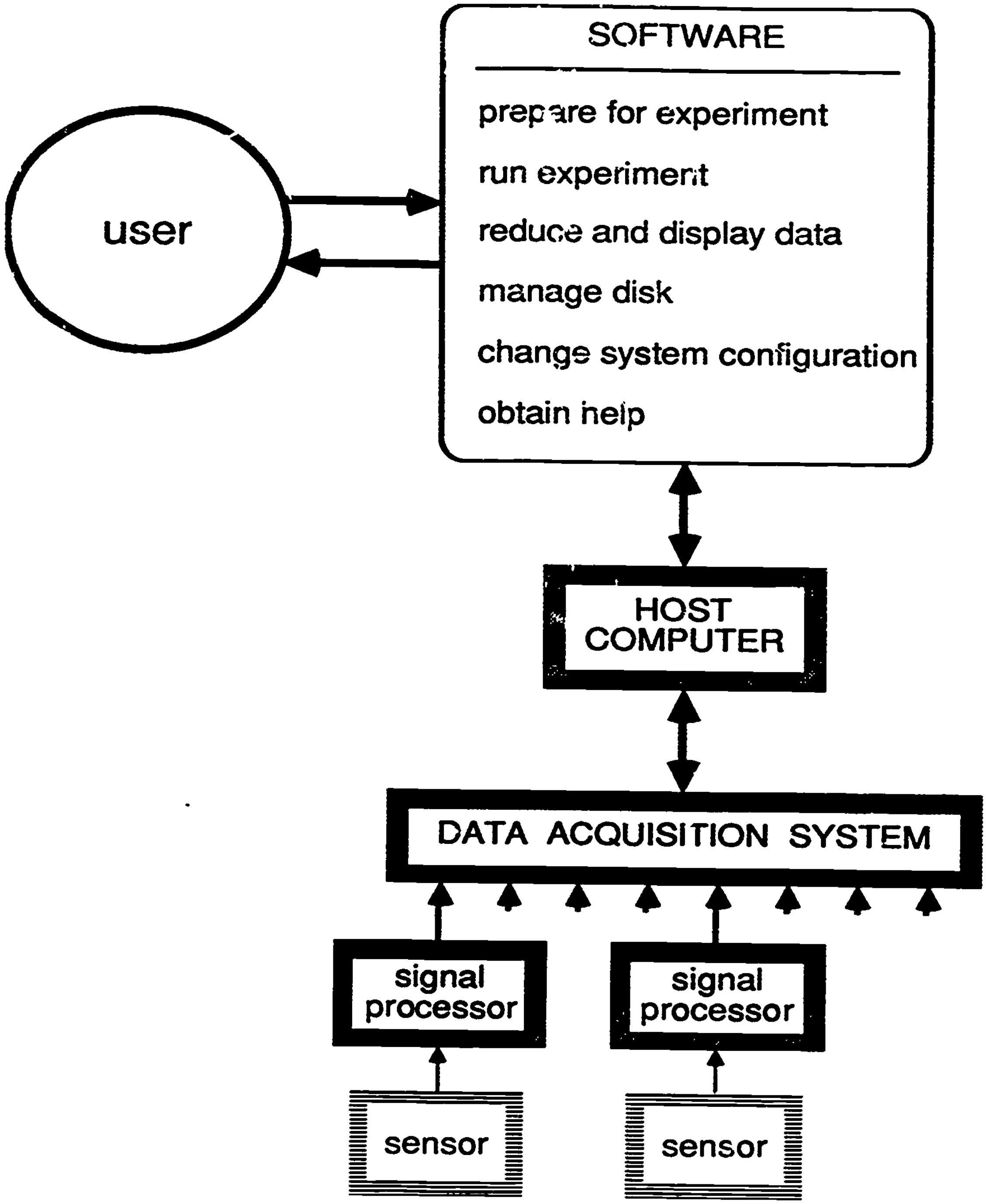


Figure I.1. A schematic representation of the design of the EDAS.

Creare

Evaluating the Effectiveness of the EDAS

We tested the EDAS's effectiveness by spending a full academic year conducting true experimental research on its actual use in the school. This year of evaluation was preceded by three months of pilot studies, conducted during the previous academic year, in which we tested the EDAS and all materials that we planned to employ in the full evaluation. During the pilot study we developed a set of desired modifications to the EDAS that were based on student and teacher feedback. Three months were then spent in refining the hardware and software to reflect these suggested improvements. During this three month period all initial data were analyzed and the evaluation procedures fine tuned.

There are two different styles in which an instrument like the EDAS can be employed in class, and we evaluated both. For some laboratories the instructor presented a demonstration of a laboratory principle for the entire class. The role of the students was to observe the teacher and answer questions during the demonstration. For other laboratories students conducted their own experiments under the teacher's guidance. Two stations were available in class, an EDAS station and a preparation station, so that the lab groups of 2 to 4 students could prepare for their experiment at one station and then move to the EDAS station to conduct the actual experimental work. In this way a full section of students could share the EDAS, with each performing all actual work on the computer.

More than four hundred students enrolled in six laboratory sciences course participated in the study. There were courses in general science, biology, chemistry, and physics. Detailed data were collected about the students' abilities and academic history from their transcripts. The evaluation of the EDAS was conducted in a series of 19 true experiments. In each experiment the EDAS was compared with the traditional laboratory apparatus used in the course. Since each of the six sciences courses met in several different sections, we were able to have the same teacher in each course present the laboratory with the EDAS for one or more sections and without the EDAS for one or more sections. To ensure that the effects we observed were not simply due to the novelty of the EDAS, we varied which sections received which set of instruments so that all students and all teachers were familiar with all methods and experienced all forms of instruction. When the EDAS was used in a section, all students in the section learned with the computer. Comparisons between the EDAS and the traditional apparatus were, therefore, made between different sections of the course. Appropriate steps were taken during the data analyses to test and control for effects of differing student abilities in the various sections of the course.

A full time research assistant and field manager was placed in the high school throughout the evaluation period. This person prepared all evaluation materials, observed student behavior in each class, entered and analyzed data, and performed all the ancillary activities associated with evaluation. In this way the teachers could proceed as they normally would have done in the courses, and disruption due to the evaluation was minimized.

We obtained information about the effectiveness of the EDAS from four sources, including semester questionnaires, daily questionnaires, class observations, and course grades:

(1) To collect information about students' long-term motivations and attitudes, we used semester questionnaires. At the beginning and end of each semester

Creare

students filled out a form that explored their attitudes, study habits, learning style, personality, school activities, and feelings about their courses and their own academic performance.

(2) To examine short-term attitudes, we had students fill out brief forms each day that a laboratory was offered, just before and just after class. These daily forms assessed students' enjoyment of the class, interest in the subject matter, the difficulty of the class, and their enjoyment of and comfort with computers. In addition, the form given before class asked about their motivation to learn and do well in the upcoming laboratory session, while the form administered after class added questions about the amount of knowledge they had learned, their confidence in what they were learning, the effectiveness of the procedures that they were using, the degree to which they worked on their own, and the degree to which they felt rushed while the class was in session. Teachers filled out a similar form with their class.

(3) To understand the EDAS impact on the actual activities of students and teachers in class, we observed student and teacher behavior during each laboratory session. During the laboratories the project field manager systematically collected continuous observations about the behavior of the students and the teacher. Every 60 seconds the observer recorded on a coded sheet which of three categories of behavior the teacher was engaged in and the proportion of students in the class that were engaged in each of six categories of student behavior. The teacher's behavior was categorized as not attending to the students, providing individual attention to students, or talking to groups of students. The students were examined for the degree to which they showed on-task or off-task behavior, that is, the degree to which they were actively and directly engaged in the laboratory assignment. The six types of student behavior were: off-task, group oriented on-task, individual engagement with the teacher, engagement with their group and the teacher, attending to the apparatus, and writing down data or notes.

(4) The overall impact of the EDAS on cognitive performance was monitored by analyzing the normal examinations and laboratory reports assigned by the teacher. For each laboratory session, data were collected about the student's actual academic performance. Tests scores were divided into scores for questions that dealt with laboratory concepts and those that dealt with other aspects of the course. Examination questions were separated according to whether they were directed towards fact recall or application skills.

Extensive multivariate analyses of all data were conducted. With 19 different experiments and the continuous attitude data, we obtained an enormous amount of information, some 200,000 numbers. The results presented here provide but a sample of the overall results, more technical descriptions of the impact of the EDAS will be made available elsewhere (Chapters II and III). For the sake of clarity and coherence, the data from one course is presented here. This course, an intermediate level chemistry course, is representative of the other courses we examined.

The Impact of the EDAS

The impact of the EDAS was manifested in three domains: the behavior during the laboratory of the students and the teacher, the students' attitudes about the laboratory, and the students' performance on test instruments after the laboratory was finished. There were differences between students who learned with the EDAS and those who did not in all three

domains. Furthermore, the impact of the EDAS was different early in the year from its impact late in the year. There were no obvious differences, however, in the impact of the EDAS across the different courses and subject matters.

Early in the academic year there were many differences between sections that used the EDAS and those in which the traditional apparatus was employed. Figure I.2 presents the observational data. The teacher's behavior was the same whether or not the EDAS was used in the laboratory. Differences in classroom behavior were particularly apparent in the observed behavior of the students. Those sections that were using the EDAS were marked by more off-task behavior among the students. There was a corresponding diminishing in the amount of on-task group behavior, attention to the apparatus, and recording of written information during the laboratory. It appears that this difference in student behavior is not due to the teacher's actions, as the teacher's activities did not vary between sections.

Late in the year the behavior of the students did not differ between the EDAS and traditional sections. Students in both sections spent most of their time on-task, engaged in working with the experimental apparatus. The initial disruption in on-task behavior with the EDAS had apparently lasted for several laboratory sessions and then disappeared once familiarity was achieved. The teacher's behavior, on the other hand, was not the same in the EDAS and traditional sections at year-end. Teachers continued to spend most of their time providing individual attention to students in the EDAS sections, but shifted to spending more time in large-group lecturing with the traditional apparatus. Apparently this effect arose from the need with traditional instruments for the teacher to give instruction to the entire class in how to accommodate to special requirements of the new pieces of traditional apparatus.

Creare

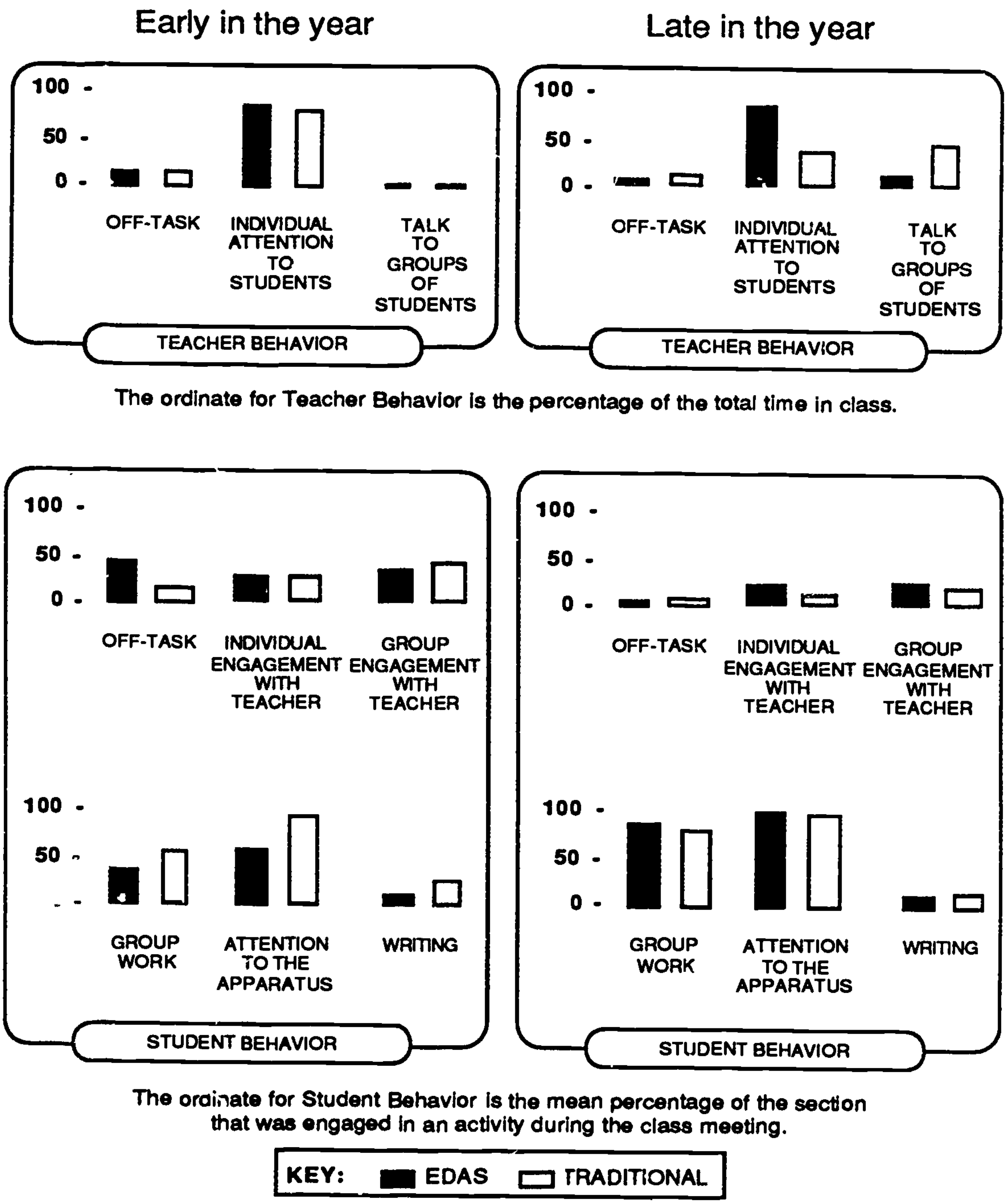


Figure I.2. The classroom behaviors of the teacher and students during typical EDAS and traditional laboratory sessions at two points in the academic year.

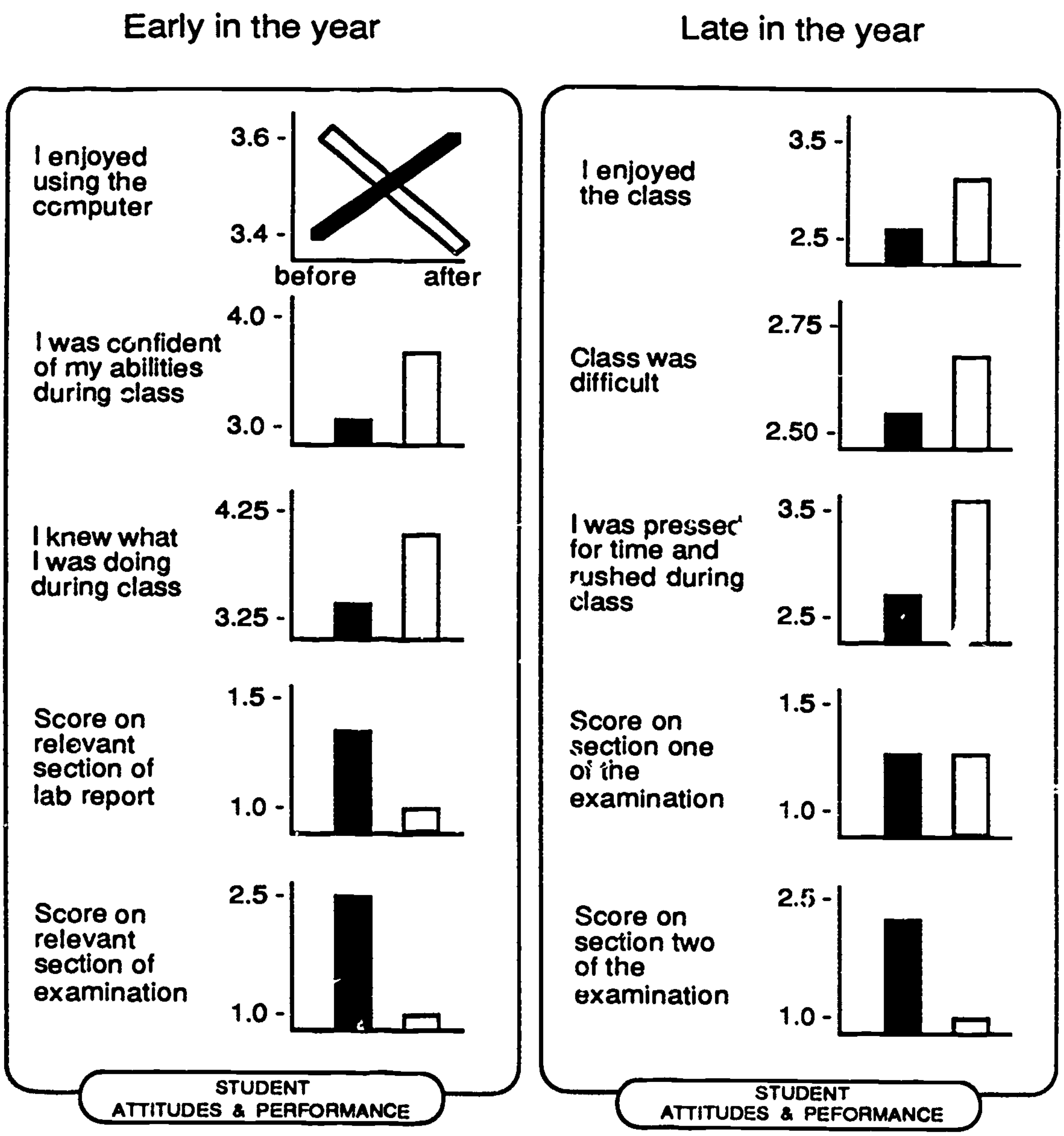
Creare

Significant differences in students' performances and students' attitudes about themselves, the course, and computers also emerged as a function of whether or not students used the EDAS. The attitudinal differences changed across the academic year. The attitudes and performance data are presented in Figure I.3.

Early in the year students who used the EDAS reported less confidence in their abilities during the laboratory and more frequent feelings of not knowing what they were doing. Despite this uncertainty, the students in the EDAS sections were more positive in their attitudes towards computers at the end of each class than were their peers in traditional sections. The actual performance of the students on the examination and laboratory reports produced outcomes opposite to the attitude effects. Students who learned with the EDAS wrote better lab reports of their work and performed better on examination questions about the laboratory exercise. Thus it seems that the EDAS students were at once performing better and challenged to perform still better.

Later in the year, the students using the EDAS in their section were equally confident and sure of what they were doing. Furthermore, they felt that the laboratory work was actually easier with the EDAS and reported that they felt less rushed than the students in the traditional sections reported. This did not lead to enhanced enjoyment of class however. Students using the traditional apparatus reported more enjoyment of class than did students using the EDAS. It is not clear how or whether this difference influences student behavior; it deserves further study. For the EDAS students the examination scores were consistently as good as or better than those for the students using traditional instruments.

Throughout the year a particular effect emerged suggesting that the improved performance on examinations for the EDAS students resulted from an increased ability to apply the laboratory principles to new problems. Rather than simply learning more facts, the EDAS students were learning how to bring what they had learned to new situations.



Note: The attitude scales range from a low of 1 to a high of 5.

KEY: EDAS TRADITIONAL

Figure I.3. Student attitude and performance comparisons for the EDAS and traditional instruments at two points during the academic year.

Creare

When the EDAS was used for demonstrations rather than hands-on laboratory experiences, its effect was different. In general, students felt more rushed when the teacher used the EDAS and felt that the use of the EDAS was less effective than the use of the traditional instruments. It is likely that this represents the teacher's awkwardness and nervousness with using the EDAS. Field observations indicated the presence of greater nervousness on the teacher's part when the computer was employed, presumably because the teacher lost the benefits of years of experience that he or she had with the traditional instruments. It is important to note, however, that examination performance did not correlate with the students' attitudes. They continued to score as well or better on examinations for which they had been prepared with the EDAS. As noted with the hands-on use of the EDAS, the improved examination performance involved an ability to apply the laboratory principles to new situations.

The Benefits and Pitfalls of Computers in the School Science Laboratory

We began the project with three questions, asking (a) whether the EDAS would operate as designed, (b) whether it would be used well by the teachers, and (c) whether it would influence the students' attitudes or performance. The field evaluation confirmed that the computer is a powerful tool that can improve science learning in crucial ways, but reminded us that it is not a panacea.

The EDAS, as a scientific instrument system, performed outstandingly well. It met all the operational design requirements we set for it. The engineering of tools is a relatively clearly understood process, and it proceeded well for us. The combination of engineers, psychologists, and teachers produced an exciting and satisfying design and development process. The team approach can provide exceptional tools to serve educational needs.

Creare

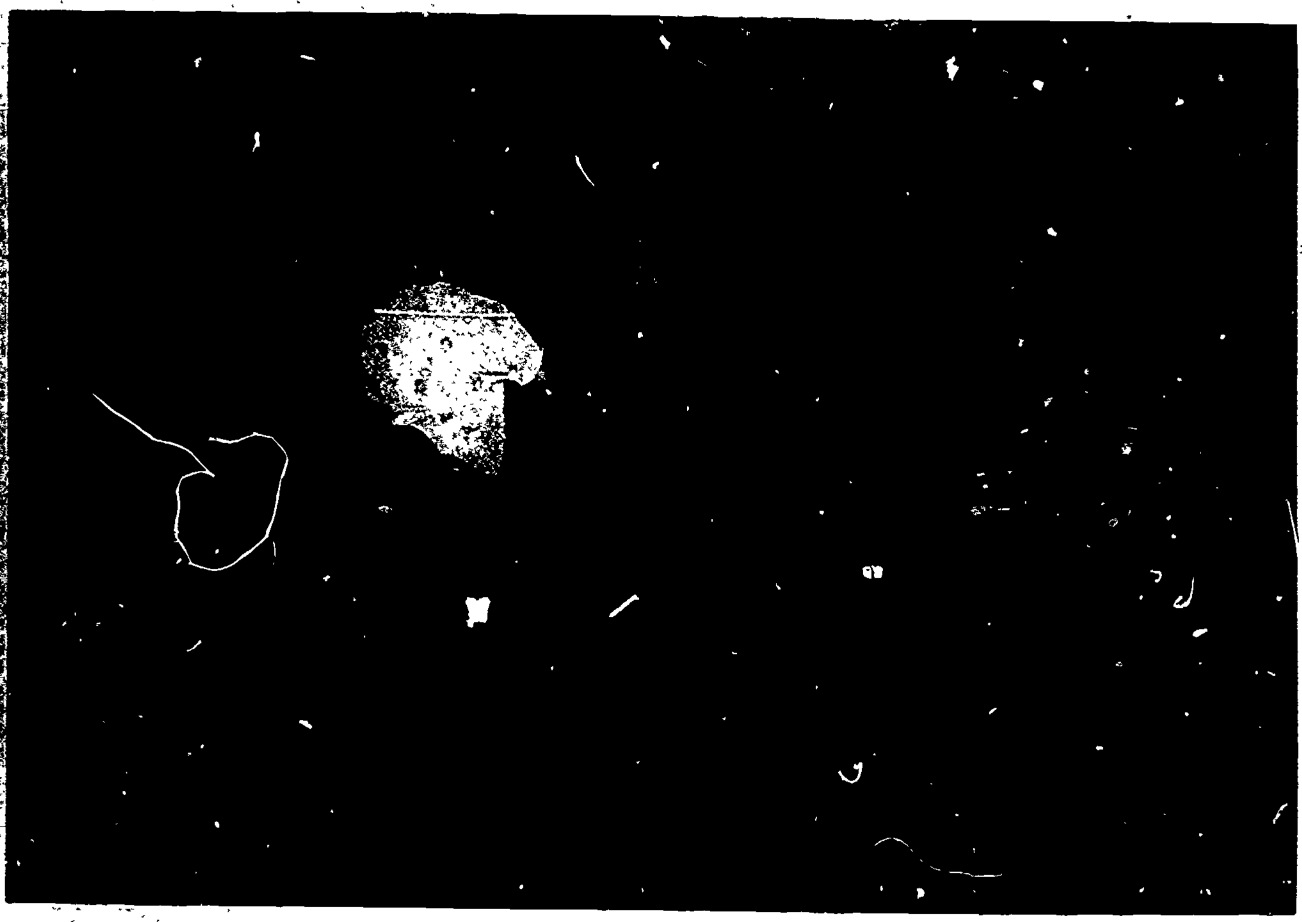


Fig re I.4. Students and teacher engaged in data collection with the EDAS. Note the degree of on-task attention from each. The EDAS is in the top right and the materials being examined in the center left of the photograph. The teacher is in the right foreground.

Creare

We found that after its development, the actual introduction of an innovative, technological tool is a complex process with a great number of variables influencing its use in the school. We encountered no absolute limits to what teachers can achieve, but we exposed problems that must continuously be dealt with, including the adjustment to a new and different tool with its concomitant changes in attitudes. Great tools need skilled users to achieve their full potential. The teachers and students worked enthusiastically to learn how to use the EDAS, yet were sometimes discouraged and confused early in their learning.

The skills needed to help the teachers with their adjustment to new technology are those of the professional trainer. For the future, training materials must be developed with the help of teachers and students who are familiar with the use of computer-based laboratory instruments. These materials should not be canned laboratory experiments; rather, they should be exercises that let teacher and student come to understand the potential for this new tool and how to operate it most effectively. The training could be brief and directed towards producing a level of comfort in dealing with the computer. It might involve a two day workshop for teachers, videotaped sessions with master teachers, or simply the presence of a motivated colleague in the school who has used the system. We noted in the test high school that teachers were attracted to using the EDAS as they heard from their colleagues and from their students of others' success with it.

The use of the computer by teachers for class demonstrations requires further study. One of the authors has used a laboratory computer to conduct class demonstrations for large groups of students for many years. It took the first few years for him to learn how to make most effective use of the computer in that mode of operation. Then the computer supported class meetings that were far more effective in impact on both student attitude and performance than those with the traditional instruments. Since the ability to ask the right question takes extended experience; one must be prepared for a learning curve with new technology and teaching methods.

In many arenas we found that while students are good at reporting most aspects of their behavior they are not particularly good at estimating what they know and whether they are proceeding along the correct learning path. Low correlations between student attitudes and actual performance appear not to be unique to the EDAS; they are commonly found in other educational research. The teacher can help students with this matter by assisting students in understanding and predicting the direction in which their classroom explorations might take them.

The presence of the EDAS also stimulated a great deal of behavior directed towards curriculum exploration and improvement. The flexibility and versatility of the EDAS enabled teachers to develop new laboratory exercizes that were simply impossible with traditional instruments. The ability to repeat experiments rapidly and with different parameters allowed students to operate in a more creative and exploratory style in class. Attention shifted from the mechanics of the laboratory procedures to the principles behind the class work. This shift is reflected in the enhanced ability to apply course principles to new areas that was found with students trained with the EDAS. The computer in skilled hands with careful thought behind its design and use can provide a powerful and effective tool to improve learning in the science laboratory.

Creare

References

Ahl, D. H. (1983). Computer-aided physics experiments. Creative Computing, October issue, 170-173.

Berliner, D. C., & Rosenshine, B. (1977). The acquisition of knowledge in the classroom. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge (pp. 375-404). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning & Verbal Behavior, 11, 671-684.

Duke, D. L. (Ed.). (1979). Classroom management. Chicago: University of Chicago Press.

Ford, B. (1983). Pixels and pendulums. Classroom Computer News, March issue, 47-55.

Harvard Graduate School of Education. (1982). Microcomputer Directory. Cambridge, MA: Author.

Horton, D. L., & Mills, C. B. (1984). Human learning and memory. In M. R. Rosenzweig and L. S. Porter (Eds.), Annual review of psychology (pp. 361-394). Palo Alto, CA: Annual Reviews.

Hursh, D. E. (1976). Personalized systems of instruction: What do the data indicate? Journal of Personalized Instruction, 1, 91-105.

Jernstedt, G. C. (1983). Computer enhanced collaborative learning: A new technology for education. T. H. E. Journal, 10, 96-105.

Jernstedt, G. C., & Chow, W. K. (1980). Lectures and textual materials as sources of information for learning. Psychological Reports, 46, 1327-1339.

Johnson, D. W., & Johnson, R. T. (1975). Learning together and alone: Cooperation, competition and individualization. Englewood Cliffs, NJ: Prentice-Hall.

Johnson, D. W., Maruyama, G., Johnson, R. T., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. Psychological Bulletin, 89, 47-62.

Kulik, J. A., Bangert, R. L., & Williams, G. W. (1983). Effects of computer-based teaching on secondary school students. Journal of Educational Psychology, 75, 19-26.

Moscovitch, M, & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. Journal of Verbal Learning and Verbal Behavior, 15, 447-458.

Piontkowski, C., & Calfee, R. (1979). Attention in the classroom. In G. A. Hale & M. Lewis (Eds.), Attention and cognitive development (pp. 297-330). New York: Plenum.

Creare

Reed, S. K., & Johnsen, J. A. (1977). Memory for problem solutions. In G. H. Bower (Ed.), The psychology of learning and motivation, Vol. 11, (pp. 161-202). New York: Academic Press.

Ryan, B. A. (1974). PSI: Keller's personalized system of instruction. Washington, D. C.: American Psychological Association.

Scandura, J. M. (1977). Problem solving: A structural/process approach with instructional implications. New York: Academic Press.

Semb, G. (1974). Personalized instruction: The effects of grading criteria and assignment length on college student test performance. Journal of Applied Behavior Analysis, 7, 61-69.

Sharan, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes and ethnic relations. Review of Educational Research, 50, 241-271.

Simon, H. A. (1980). Problem solving and education. In D. T. Tuma and F. Reif (Eds.), Problem solving and education (pp. 80-96). Hillsdale, NJ: Erlbaum.

Smith, K., Johnson, D. W., & Johnson, R. T. (1981). Can conflict be constructive? Controversy versus concurrence seeking in learning groups. Journal of Educational Psychology, 73, 651-663.

Spitzer, H. F. (1939). Studies in retention. Journal of Educational Psychology, 30, 641-656.

Webb, N. M. (1983). Group composition, group interaction, and achievement in cooperative small groups. Journal of Educational Psychology, 74, 475-484.

Welch, W. W. (1979). Twenty years of science curriculum develoment: A look back. In D. C. Berliner (Ed.), Review of Research in Education, No. 7 (pp. 282-308). Washington, D. C.: American Educational Research Association.

Wittrock, M. C., & Lumsdaine, A. A. (1977). Instructional psychology. In M. R. Rosenzweig and L. S. Porter (Eds.), Annual review of psychology (pp. 417-459). Palo Alto, CA: Annual Reviews.

Creare

CHAPTER II: THE METHODOLOGY OF THE PROJECT

Evaluating The Impact of Computer-Based Instruction in the Schools

by G. Christian Jernstedt
and
Tyrone D. Cannon

The claims about the impact of computer-based instruction (CBI) on students range from the elimination of schools as we know them today (Papert, 1980) to the creation of a culture of psychopaths (Sardello, 1984). The debate about the impact of CBI has focused on both the style in which computers should be used in classrooms and the manner in which the outcome from CBI should be measured (Cumming, 1984).

That computer-based instruction can lead to significant improvement in learning has been clearly determined (Kulik, Bangert, & Williams, 1983; Kulik, Kulik, & Cohen, 1980). However, the enormous variety of CBI products available for instructors and the rich diversity in the actual use of computers in the classroom have made direct, useful applications of research results difficult to accomplish. Evaluation of the actual, in-use effectiveness of CBI materials and methods must, therefore, be a regular part of employing these tools in the classroom. Such evaluations can be called "applied" in that they occur within the context of the classroom where pedagogical priorities often must take precedence over evaluation needs. For example, random assignment of students to classes may not be possible in a school where tracking of students by ability is a policy. The need for continuous, integral, valid evaluation of CBI tools and styles of use is critical if they are to be understood and used optimally. Unfortunately, evaluation, and especially applied evaluation, is a complex enterprise and few instructors routinely conduct such evaluations of their course materials.

We examined current evaluation techniques and designed a process by which continuous, integral, and valid applied evaluations could become a part of the use of CBI technology. The actual project centered around the design and testing of a new computer-based educational data acquisition system (EDAS) for school science laboratories. The intent of the project was to compare the effectiveness of the EDAS with that of traditional apparatus in the performance of laboratory science exercises at the secondary school level. The evaluation took

place at a regional high school during the full academic year, 1985-1986. We worked in conjunction with the developers of the EDAS, as well as the teachers and administrators of the school, during both the planning and implementation stages of the evaluation. We will use the evaluation activities of this actual project as an example around which we develop the central ideas of how CBI evaluations may be conducted in school settings. Using this actual example, a case study, not only provides concrete illustrations of the general points being made, but also ensures that the problems and pitfalls of applying theory to practice will be revealed and considered. The methodology of applied CBI evaluation is the focus in the current article; the data from the particular study by which the general evaluation procedures were tested are reported elsewhere (Jernstedt, in press; Jernstedt & Rothe, 1986).

THE BASIC EVALUATION DECISIONS

Decision Making Resources

There are many written resources that can be of substantial help in conducting an applied evaluation. Judd and Kenny (1981) provide the best and most appropriate discussion of the basic issues in such evaluations, though technical sophistication is needed to follow the book. For many educators with some experience in evaluation, the handbooks by Rossi, Freeman, and Wright (1979) and by Isaac and Michaels (1971) provide encyclopedic reference information. The classic resource for applied evaluations is by Cook and Campbell (1979). A broad, philosophical, yet directly applicable resource is Cooley and Lohnes's (1976) description of educational evaluation. Throughout the present article we summarize and expand on ideas presented by these researchers.

The Goals and the Type of Evaluation

The most basic question we had to ask was why the evaluation was being undertaken? We identified four reasons for conducting such applied CBI evaluations: (a) to obtain knowledge of general interest about the learning process and CBI materials, (b) to provide a measure of accountability to educators and administrators who purchase or use CBI materials, (c) to guide future improvements in the particular program we were studying, and (d) to obtain funding to support further program development. As is typically the case, we found elements of all four goals in the project. Principally, however, we were interested in understanding the program and describing its possible effectiveness. We thus were employing both of the two types of evaluations: formative and summative. Our evaluation was formative in that it was directed towards feedback of information about the program while the program was occurring. The evaluation was also summative in that part of our focus was on specifying the outcomes from the CBI program in a general sense that could be employed by others in other areas besides science laboratories. We also found that our evaluation was both a process and an outcome evaluation, concerned with understanding the processes that were producing the effects of the CBI as well as with uncovering the effects themselves. Where we did not find a blending of different styles was with the perspective of the study. Applied evaluations of CBI programs tend not to blend the two different theoretical perspectives of exploration and hypothesis testing. Applied evaluations typically are exploratory in nature, seeking to find all possible effects of the CBI intervention, rather than to test a particular hypothesis.

It became clear to us that although one can describe in a theoretical sense a hierarchy of the various types and styles of evaluations (Venezky, 1983), in practice elements of most styles

Creare

and types tend to be present. The issue thus becomes one of making explicit the many implicit decisions embedded in designing the actual procedures. We found that evaluation is best seen as "system learning," in the sense that it is a process by which a set of individuals in an organization learn about the complex interaction between their own behavior, the functioning of the materials and resources they use, and the behavior of their learners.

The school is a rich source of information; as a natural environment it affords researchers the opportunity to evaluate programs and innovations in the same context in which those programs are to be implemented and used. Results obtained in such a setting are well suited for application in other, similar environments. While the school provides an advantage to researchers in terms of the generalizability of findings, it is also a difficult setting in which to maintain control over experimental procedures. As in other naturalistic settings, many forces other than the experimenter can affect the manipulation. The present study provides an example of how the structure of experimental design can be superimposed on the school setting, thereby preserving the field environment and maintaining the control necessary for rigorous scientific research. We remain convinced, as we were when we began, that skilled educators can conduct high quality, valid evaluations as they educate. They may require additional assistants or funds, depending on the size of the evaluation, but good evaluation and good education can occur simultaneously. Furthermore, we believe that the two must occur together if the educational process is to maintain or increase its quality and effectiveness.

MEASUREMENT ISSUES

Variables to be Studied

The choice of what variables to sample determines in large part what results can be obtained. In an applied evaluation we maintain that one must cast a wide net in choosing variables. We chose variables from all three of the domains that influence educational outcomes: Person, Behavior, and Environment (Bandura, 1978). From the Person domain, those characteristics that the student brings to the classroom, we sampled the abilities, previous school experiences, personality, and attitudes of the students. In the Behavior domain, the actual activities in which the learner engages, we directly observed the students during class meetings and obtained reports from them about what they did when not in class. In the Environment domain, the characteristics of the learning environment itself, we recorded the attributes of each course in which the CBI materials were used.

Though it might be desirable to assess students using as many different techniques as possible, such a procedure can overwhelm the evaluator with too great a variety of information. We obtained data with most of the different assessment techniques that are available, but concentrated our efforts on four: demographic records in the school, student self-reports, behavioral observations, and teacher created performance measures. We occasionally used interviews with students and teachers to collect a supplementary set of field notes to help flesh out the quantitative data during the later reporting stage, but decided that interviews could not provide the objective validity that the CBI materials required. For similar reasons, the student self-report measures were obtained with well-researched, published and locally developed questionnaires and did not include open-ended questions for extended student written comments. Furthermore, with the natural availability of teacher prepared objective performance measures we did not need to employ reports from significant others in the students lives to obtain verification of their self-reports. We were especially careful to choose

variables that revealed demographic, cognitive, and affective aspects of the students. Since there are often differences between preferences for and achievement with CBI (Burger, 1985), the distinction between cognitive and affective was maintained with the outcome measures as well; we collected information about both how the students performed and how they felt they performed.

The Apparatus and CBI Resources

The computer-based system used in the study, the Educational Data Acquisition System (EDAS), was developed by a team at Creare, Inc. composed of engineers, psychologists, and educators with funding from the U. S. Department of Education. The EDAS is designed according to sound human factors standards (Sawyer, 1985) and consists of several independently developed components and a package of software and documentation that integrates the components. The hardware includes an IBM PC with dual drives, serial interface, printer, and monochrome monitor; a Burr-Brown Data Acquisition System with an analog to digital converter, microprocessor, RAM memory, and various signal processors; and a set of standard sensors from a variety of manufacturers that are capable of detecting pH, temperature, pressure, motion, force, conductivity, sound, light, and electricity. The pedagogical core of the EDAS was the software. A comprehensive software package was written to allow students to use the EDAS in much the same way that a university or industry research scientist would use a computer. In actual use, the student scientist with the EDAS makes choices through a series of menus that provide for the collection, storage, analysis, and display of data obtained by the sensors.

The design of the apparatus and software highlights a philosophical perspective that guided decision making throughout the project. A dilemma arises whenever one has to make a decision involving the comparisons to be made in the evaluation between the experimental treatments being designed and tested and the control treatments (the status quo). Should one choose the best possible experimental treatment or test a range of possible treatments? Should one test a modest experimental treatment that is likely to be available to the schools or one that is exceptional but is unlikely to be available? We maintain that in an applied evaluation there is an optimal answer to these questions: one should test the best possible treatment that is likely to be used by the schools. For example, we were faced with deciding how many computer stations there should be in the laboratories. Optimally each student or pair of students should have their own computer. But this will rarely be the case in typical schools. Thus we limited ourselves to one computer station per classroom. On the other hand, when deciding whether the computer system should be a powerful one with large memory and other such features, we chose the powerful model. Our reasoning was that we wanted to know how a good computer would affect science learning not how a rudimentary computer would affect the learning. We decided that subsequent studies would be the appropriate place to examine the degree to which the power of the machine influences any effects that we might observe. The first study in an applied evaluation, we reasoned, must examine the maximum contrast possible between the new experimental treatment and the traditional apparatus. Thus we recommend that, when making decisions about the experimental treatment, applied evaluators adopt the perspective of choosing the maximum contrast within the limits of what can actually be accomplished in the natural classroom; in short, create the best, but employ it realistically.

Creare

Research Setting, Subjects, and Sampling Techniques

The setting of the study was a medium-sized, regional public high school located in Hanover, New Hampshire. The high school was chosen because (a) its students were representative of many other small, good high schools; (b) an agreement could be arranged with the administration and faculty to have all science courses use both the EDAS and their traditional laboratory exercises, (c) full data could be obtained about students, faculty, and courses, and (d) the school was near the project headquarters. The population of Hanover and the surrounding community is largely associated with Dartmouth College, an army research laboratory, and the regional medical center. The high school, however, draws from a much larger area, representing both towns and small rural communities.

The high school includes students from grades nine through twelve. Since the project was concerned with CBI in science learning, only the science department was involved in the study. The high school science department offers six courses in the four traditionally taught science areas: 1) a general introductory course in physical science; 2) two courses in chemistry, one general and one advanced; 3) two courses in biology, one general and one advanced; and 4) one course in physics with three levels of difficulty. A regular course sequence of increasing difficulty is followed by most students. All science courses are two semesters in duration. There are six faculty in the department, approximately one for each course.

Only students who enrolled in one or more of these six courses participated in the study. The total sample size was 461 (209 females, 252 males). Although the mean ability level of the sample is above average, the range in standardized test scores is in keeping with national norms. Student participation was considered voluntary in principle; any student could refuse to take part. However, because the study took place in regularly scheduled classes and involved the regularly planned learning materials, no students declined participation. Elaborate measures were taken to protect the confidentiality of the data so that the evaluators could collect detailed information about all aspects of the participants' lives.

The Relationship of the Evaluation to the School

The evaluation was designed to preserve the natural environment of the classroom and, at the same time, to superimpose the structure of experimental control over classroom proceedings. To achieve this synthesis, we met with the science teachers extensively and continually throughout the project. The curriculum of each course was examined, and laboratories which met three criteria were chosen for experimental comparisons: (a) the EDAS could sense the process in question, (b) the variables involved could be graphed against time, and (c) the system could be used in an analogous way to the traditional apparatus in demonstrating the process or principle in question.

The laboratories chosen for experiments were taken from the regular curriculum of the courses; no new laboratories were developed. However, there was a significant amount of activity associated with pilot-testing labs, which did fall under the rubric of curriculum development. Much of this work involved adjusting the materials of the textbook laboratory to conform to the specifications of the EDAS and its sensors. Often the quantities of substances had to be altered in order to fall within the sensitivity range of the sensors. In other instances, the substances themselves were changed so that an EDAS sensor could be employed. These

altered laboratories, however, retained the basic procedures and conveyed the same principles as the established textbook versions. In addition, any changes in materials that were made to accommodate the EDAS also applied to the traditional treatment groups.

Since part of the evaluation strategy was to employ the EDAS as realistically as possible, we examined two different styles of use. For about half of our experiments we studied the use of the computer to assist the teacher in presenting a demonstration to the class. In the other half of the experiments we examined the direct use of the computer by students in laboratory sessions. This 50/50 proportion approximated the actual proportion that the two styles of use were employed by the teachers.

In applied evaluations it is not always possible to obtain uniform usage of the experimental treatment across all possible conditions in the school. Some teachers were more willing to participate in the present study than others. In general, it was noted that the teachers initially appeared resistant to change in their teaching practices and reluctant to take initiative in finding applications of the EDAS in their own curricula. This observation is supported by recent studies which demonstrate considerable resistance to technological innovations on the part of teachers (Stimmel, Connor, McCaskill, & Durrett; 1981). Some of the teachers' reluctance to participate may also have been due to the additional time and effort required to pilot-test laboratories for the EDAS.

We counteracted this situation by maintaining a constant presence in the school; providing the initiative to plan, pilot-test, and schedule labs; assuming all or most of the work associated with setting up and testing the EDAS for each lab; verbally reinforcing with the teachers the need to conduct many experiments in each course; and offering several forms of compensation to both the teachers and the school. The importance of this presence cannot be overemphasized for applied evaluations. A semester of work in the year prior to the actual evaluation yielded little information when we did not maintain this presence in the school. One might question whether the presence actually changed the classrooms enough to distort the results. Although no concrete evidence is available to answer this question, both the teachers and the students reported that the presence and activities of the field manager encouraged and enabled them to participate in experiments that they would not have otherwise done, but did not actually change their behavior during the experiments. We noted that teachers' willingness to participate in the project improved after first use and continued to improve over time, as they found the work load to be approximately what it normally was for them.

The teachers' initial reluctance to involve themselves significantly in the project threatened to reduce the total number of experiments and the scope of the evaluation. Although this problem was fortunately resolved, we were also concerned with the potentially disruptive effects of teacher negativism or enthusiasm toward the EDAS. We counteracted the possibility of deleterious instructor effects by carefully outlining the evaluation's experimental procedures and indicating the potential sources of systematic error to the teacher before each lab. In addition, if the laboratory was a teacher demonstration, we observed while the teacher practiced with the procedures using both the EDAS and traditional apparatus during pilot-testing, and we indicated those aspects of the teachers' delivery which needed to be changed in order to make the presentations equivalent. Finally, we also maintained a presence in the actual classroom as observers; we discarded the data from labs in which extraneous events caused the classroom proceedings to differ between treatments.

One must exercise great care in dropping data. Since the goal is to obtain an accurate picture of what is occurring, when problems during an observed class are likely to provide a false view of events, those data must be dropped. For example, if during a class session the traditional or CBI apparatus failed, a fire drill or other such unusual interruption occurred, or critical comparative data were missing, then one must not included data from that session in the analyses. On the other hand, it is crucial that data not be discarded because it appears that the results are not in the expected direction. The discarding of data or experiments must occur prior to the determination of the relative outcomes from the experiment or serious bias will destroy the validity of the evalution.

We often found ourselves presuming a degree of structure, organization, and adherence to schedule that was not evident in the school. Some teachers were more organized than others, and for some courses it was possible to plan all of the experiments for the year at the beginning of the project. For other courses, the field manager checked in with the teachers at monthly intervals to plan laboratories for each time period.

The planning aspect of each laboratory consisted of obtaining and connecting the necessary sensors, pilot-testing the actual laboratory procedures with both the experimental and traditional apparatus, designing the experimental procedures by which to compare the two instructional techniques, developing lab manuals and instructions where appropriate, and scheduling the experiment. This planning was directed by the teacher with the full support of the evaluation field manager. Although some teachers rigorously followed their planned curricula, most courses were significantly behind schedule in their syllabi at the end of the first semester and at the end of the year. The teachers often cancelled laboratories and complete sections of the course curricula in response to these delays.

It was recognized, however, that the presence of the researchers and the project itself were secondary in importance to the presence of the teachers and students and the enterprise of education. In all of our relations with the high school, we stressed our understanding of and agreement with this hierarchy. At the same time, we emphasized the need for multiple tests of the EDAS and for rigor and precision in experimental procedures. In practice, there was a considerable degree of cooperation and flexibility maintained by both the teachers and experimenters.

Compensation to the school

We noted previously that various forms of compensation were offered to the teachers and the school in exchange for their participation in the study. Compensation was an essential feature of the evaluation, since we required use of both the personal resources of the teachers and students and the physical plant of the high school. Although we initially expected to pay the teachers directly for extra time they spent that would not have been required in the normal course of events, that became unnecessary. The presence of the field manager eliminated most of this extra time requirment for the teachers. Compensation to individual teachers took the form of increased assistance for the teachers from the field manager. The form of compensation for the school involved the transfer of ownership of the EDAS to the school upon completion of the project.

In addition, we were able to offer the science department compensation in the form of a wealth of information about its courses. For each course we provided data on the effectiveness

Creare

of the learning materials, the instructor, and the various instructional media, on student perceptions of and attitudes toward the course, and on chara‹ ›ristics of students who enrolled in the course. These data were obtained from surveys admin‹ ›ed to all science students three times during the year.

The school's experiencing of the evaluation, itself, also produced other indirect gains: (a) considerable curriculum development, planning, and structuring was associated with the evaluation; (b) data on the relative merits of instructional approaches were collected and analyzed; and (c) the practical experience with evaluation research prepared the school and the students for subsequent investigations both in-house and by outside evaluators. All of these features have survived beyond the term of the project. The science department has included in its own budget for 1986-1987 funding for the administration and analysis of the surveys we constructed for the present evaluation.

THE RESEARCH STRATEGY

As indicated above, the principal strategy of this applied evaluation was to mesh the research design and procedure with the curriculum of the courses and the normal functioning of the classrooms. In the following discussion, we describe the methodology of the study and the tactics employed in executing this general strategy.

Experimental Design and Group Assignments

We applied a relatively novel approach to applied evaluation by designing the study to be analyzed with meta-analytical techniques (Glass, 1976; Wolf, 1986). We conducted nineteen individual experiments in which the EDAS was compared to traditional laboratory science instruction. Computations of effect sizes, in additional to statistical significances, were employed to survey the overall effectiveness of computer-based science teaching and to assay the influence of various study features on both attitudinal and educational outcomes. These analyses are innovative for applied evaluations in that they are applied to data obtained in the same setting, using the same subjects, and employing standard instrumentation, design, and procedures. We highly recommend them. These procedures offer the ability to make broad yet accurate summaries of a huge body of data. The conclusions can be stated more clearly and simply than can the results of 19 separate experiments. Finally, the meta-analytic, quantitative summaries are more likely to generalize accurately that would a verbal summary of the 19 different sets of results.

In an applied evaluation the choices among the three factors that differentiate between research designs (Judd & Kenny, 1981) are fairly obvious. The first factor, how learners are assigned to the different treatments to be tested, is typically determined by the existing class arrangement. The second factor, whether a pre-treatment measure of the subjects is taken, simply must be a part of an applied evaluation because of the many forces preventing random assignment of students to treatments. The third factor, whether all participants experience all treatments or only some, is best handled by repeating the treatments enough times that all participants do receive all treatments. The meta-analytic procedures described above are quite consistent with this decision.

For the present study, two treatment groups were established for each experiment, one using traditional apparatus and procedures and one using the computer-based EDAS approach

Creare

in the performance of laboratory exercises. The sectional divisions of each course were followed in forming group assignments for the 19 experiments. We randomly assigned use of the computer to one to three sections and use of traditional apparatus to one to three sections, with the total number of treatment groups dependent on the number of sections in each course. In some experiments, however, the length of the exercise prohibited use of the computer for more than one lab group (4 to 6 students) within each section; in these instances, we randomly assigned use of the EDAS to lab groups. This is an example of turning a prescribed condition in an applied evaluation into a design feature. By so varying the manner in which lab groups accessed the computer, we were able to make an additional, unplanned comparision, between two different, realistic ways of employing the computer in class.

Because most students follow a regular course sequence in the science curriculum, class year was a major determinant of enrollment for each course. In addition, ability level contributed to course selection and to sectional placements within a course. Although the science courses lacked a sufficient number of ability-selected sections to create a higher-order design (treatment by ability), we included steps in the analyses to check for the potential effects of ability on the treatment outcomes and control for ability as a possible confound through analysis of covariance.

Lab groupings were teacher-selected at the beginning of the year. Although method of grouping has been shown to affect performance in and attitudes toward laboratory science courses, we did not attempt to manipulate this variable in the present evaluation. The teachers altered lab group assignments on some occasions throughout the year, but these decisions were made without respect to the evaluation. We made no changes in the established groupings for experimental purposes.

Student Data Measures

Our goal in collecting data was to use standard, validated questionnaires and tests wherever possible. Table II.1 presents an overview of the data that were obtained about the students. The standardized test scores, grades and demographic profiles of the subject population were obtained from the transcripts on file at the high school. The standardized tests included the Metropolitan Achievement Test, PSAT, SAT Aptitude and Achievement Tests, and a national test administered only to sophomores. Not all students took all tests. Throughout the data analyses we never estimated missing scores, instead for any given comparision we employed only thoses students for whom we had full data.

Grades from all previous course work in English, science, and mathematics were obtained for each student. No achievement data were available for the current year, except for the students' semester and final grades in their science courses. This situation meant that the only achievement data available for freshman subjects were current science course grades.

Demographic information collected for the evaluation included class year, year of birth, birth order, family size, gender, junior high school, and number of years in attendance at the high school.

Creare

Table II.1
Variables and Data Measures

Domain	Source of the Data	Measure Obtained
Person	school transcript	standardized test scores courses taken previous grades
	self-report inventory	demographic information personality study habits learning style attitudes towards school
	daily class questionniare	attitudes towards class attitudes towards computers
Behavior	observational checklist	on- and off-task activities
	interviews	actions during class
Environment	school records	course characteristics
	self-report inventory	instructor characteristics course characteristics
	interviews	instructor characteristics course characteristics
Outcome	teacher graded work	lab reports quiz scores examination scores

Creare

Based on previous psychometric work and extensive pilot studies, we constructed a self-report inventory using 5-point scales to assess students' learning style, study habits, academic and social self-esteem, and satisfaction with their school performance. A number of the personality questions were directed towards self-efficacy, since perceived self-efficacy may be an important aspect of school achievement (Schunk, 1984). In addition, the inventory asked students to evaluate various aspects of their course, including instructor characteristics, grading criteria, and overall quality. The original version of this survey was composed of 160 questions; this form was administered during the pilot-testing phase prior to the actual evaluation and factored to the current form, which consists of 69 questions. Three versions of the inventory were established, all of which were equivalent in form and questions asked, but which differed in terms of length and timing. One version asked students to answer according to what they expected their science course to be like for the current year; one asked students to answer according to what they had experienced in their science course during the first semester; and one asked students to answer according to what they had experienced in the science course during the second semester. The last two versions included 41 additional questions which asked students to rate the various instructional media and course materials on three dimensions: usefulness, difficulty, and integration with the rest of the course. The pilot phase of the project before the actual year of evaluation began was of major value. We were able to local many sources of difficulty and improve many procedures before they could jeopardize the actual evaluation.

Treatment Outcome Measures

A daily report questionnaire was constructed using 5-point scales to assess pre- and post-class attitudes in the experimental lab sessions. On these daily reports students were asked to rate their enjoyment of the class, their interest in the subject matter, the difficulty of the class, their enjoyment of using computers, and their level of comfort in using computers. In addition, two unique questions were asked on the pre-class form to assess before-treatment motivation to learn and to do well. Six such unique questions were asked on the post-class survey: the amount of knowledge learned in that day's class; students' confidence in their abilities; the effectiveness of the methods used to learn; the students' certainty with the procedures; the originality of their work; and the hurriedness of that class period.

An observational checklist for the assessment of teacher and student behaviors was developed with extensive pilot work and used in the class labs. Three categories of teacher behaviors were included in the checklist: off-task to all students (e.g., out of the room); providing individual attention; and providing group attention to all students (e.g., lecturing, as in a demonstration). Observations of teacher behaviors were recorded at one-minute intervals. The three categories of teacher behaviors were considered mutually exclusive; only the single category most displayed during an observational period was recorded. Six types of student behavior were included in the checklist: off-task; group-oriented (in the performance of the task); individually engaged with the instructor; attending in a group to the instructor; attending to the apparatus; and note-taking. Estimates of the proportion of the class engaged in each category of behavior were recorded after each successive minute of class time. Types of student behavior were not considered unique; the same student who displayed two types of behavior during a single interval was counted in the total proportion for each category. The proportions of the class engaged in all categories in an observational period could sum to more than one. Five fractions of class participation were used by the observers: 0, representing no students; 1/4, representing from 1 student to 1/4 of the sample; 1/2, representing more than 1/4

Creare

of the sample but not more than 1/2; 3/4, representing more than 1/2 of the sample but not more than 3/4; and 1, representing more than 3/4 of the sample.

Assessment of academic performance was made by obtaining teacher-graded reports, quizzes, and exams. No special performance measures were constructed; only the the assignments included in the normal grading process were analyzed. Data prepared for analysis included scores on each question of a particular exercise that related to the students' work during the laboratory; subscale scores formed by summing the raw scores of all related questions; and the total scores of each exercise. The subscale groupings were created to classify student performance domains: fact recall, use of principles, and application or extrapolation of information to new situations. Variables such as these subscale groupings have been labeled measures of scientific literacy (Arons, 1984).

Procedures

The field manager served as experimenter throughout the evaluation. Not only did this give him full control of all procedures and ensure the maintenance of good scientific methods, it also encouraged the students to assume a more careful manner when filling out forms, as his presence came to represent for them a thoughtful self-reporting of their states.

The sequence of procedures followed during the evaluation is illustrated in Figure II.1. The self-report inventories were administered in each course during the first two weeks of the first semester, in the first two weeks of the second semester, and in the final two weeks of the academic year. Administrations were made in individual class meetings. For all administrations, the experimenter read standard instructions, part of which cited examples of how student responses were taken seriously and were making a difference in the way the courses were taught and organized. These statements to the students have a positive effect and can significantly improve the accuracy and validity of the self-report measures.

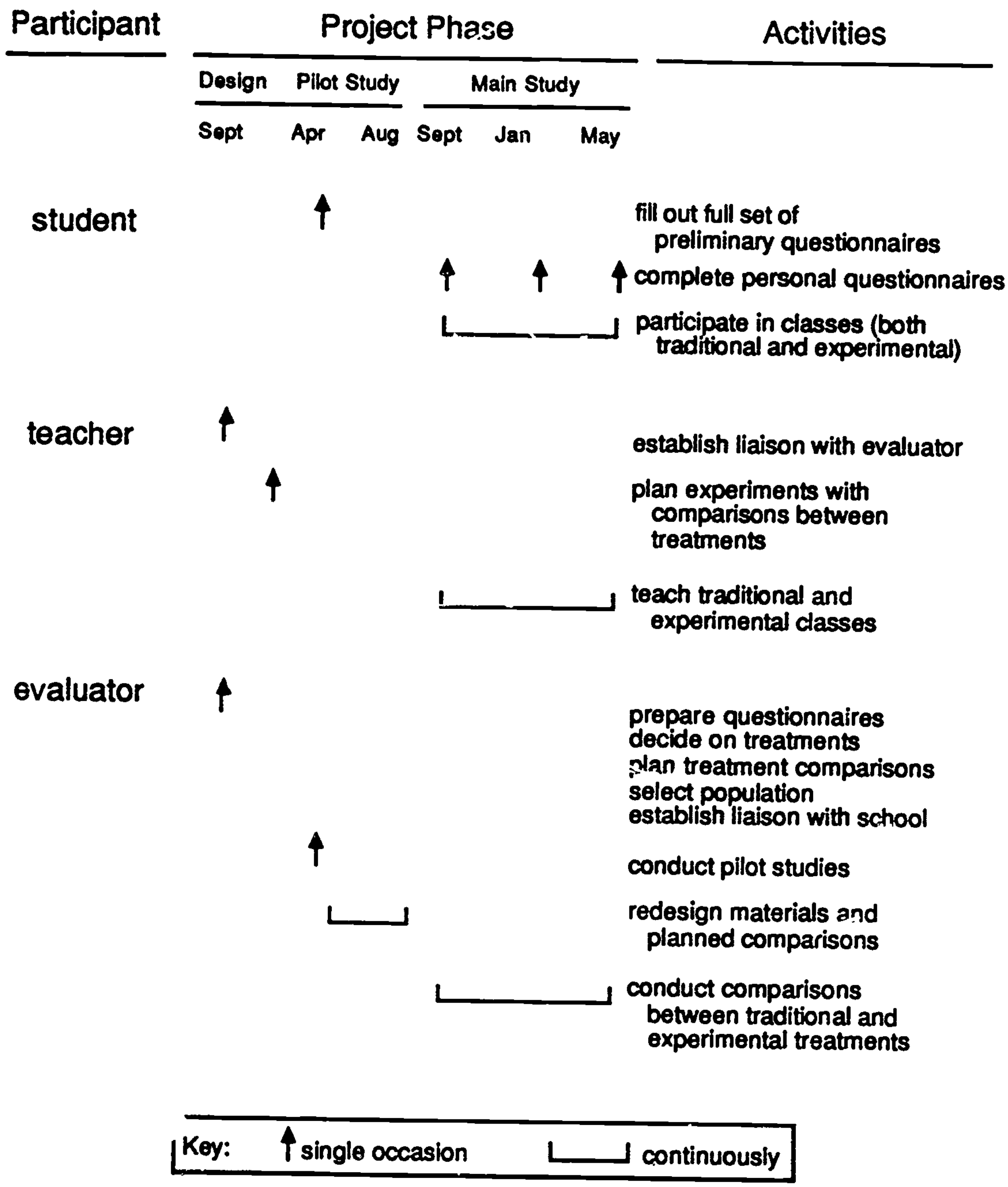


Figure II.1. The temporal sequence of activites engaged in by each of the particioants in the applied evaluation.

Creare

Experimental comparisons of the two laboratory instructional techniques were conducted throughout the academic year. These experimental labs followed the regular topical sequences and schedules as established in the course syllabi. A standard procedure was followed in each experiment. After attendance was taken by the teacher, the experimenter addressed the class, indicating that as part of the on-going evaluation of their science course, the students were being asked to fill out two questionnaires in that day's class, one at the beginning and one at the end. To prevent the experimenter's presence from being strongly associated with the EDAS, so that students expected to use the system in particular sessions, the experimenter also evaluated class meetings in which experimental comparisons were not being taken. In order to prevent expectancy effects from distorting the pre-class daily reports, the EDAS was present in the room during the introductions for all sections and in plain view of all students. Once all students were finished filling out the pre-class report, the experimenter assigned lab groups to treatments when appropriate, and then proceeded to record observations of teacher and student behaviors. In those experiments for which sectional group assignments were made, the observers filled out checklists for each section. In those experiments for which treatment divisions were made by lab groups, the observers alternated recording one-minute interval observations between the two treatment groups within each section. In all studies, when the first lab group finished with the laboratory procedures, observations were terminated and the post-class questionnaires were handed out to each group of students as they finished.

The dual role of the field manager as experimenter and guide for the teachers in the use of the EDAS posed a problem. On some occasions the apparatus developed technical problems while the students were performing the laboratory exercises. In most cases the students or the teacher could correct the problem. For major problems, however, the experimenter was the only person present with sufficient knowledge of the system to attend to them. If these problems were easily and quickly corrected, the data from that section were retained; however, if the problem necessitated prolonged and involved attention by the experimenter, the data from that particular section were discarded. Though these problems were infrequent, having a protocol prepared to deal with them was important for minimizing interference with both the teacher's and the evaluator's goals.

Data Processing

The preparation of data for analysis involved converting each raw set of data into a standard matrix format. We employed a variety of multivariate data analysis procedures to explore the data, using a package of software that provided a high level of interaction between investigator and data. It is very easy to lose sight of the data in a complex evaluation due the sheer number of numbers. The use of interactive statistical packages maintains a closer contact with the numbers being manipulated. Our policy was to take the conservative approach. Where multiple tests were made, we adjusted significance levels to reflect the many tests being performed. We tested all assumptions about distributions of data that were required for the various statistical tests. We graphed data frequently to directly observe the shape of effects we were studying. We strongly urge a similar approach for other investigators. A large data base and powerful analysis techniques are seductive to the evaluator; conservative procedures can prevent analysis from becoming a fishing expedition for results. The goal must always be that of discovering the patterns that exist in the data, not in proving certain hoped for findings.

We expected that we would need to covary academic ability in many of our analyses because our outcome measures involved traditional lab reports and examinations and because

Creare

some classes were stratified by ability. Our intent in collecting the many different standardized tests was to find an objective measure of ability that could be used as a covariate in the analysis of the treatment outcomes. To this end we created a single ability matrix by combining all aptitude and achievement subscales in a single matrix. This matrix was then factor analyzed. The resulting rotated factor pattern revealed a three-tiered hierarchy, accounting for 70% of the total variance. The factors were (a) general mental ability, which accounted for 50% of the variance; (b) demonstrated achievement (15% of variance), best represented by cumulative class rank; and (c) specific mental ability (5% of variance), in the area of mathematical abstraction. We chose the semester grades in the science course as the best representative of the general mental ability factor and used it as covariate in the analyses.

With the observational checklists for each category of student behavior, we counted the total number of times each of the five possible fractions of class participation (0, 1/4, 1/2, 3/4, 1) occurred. We also counted the frequencies of each category of teacher behavior. This procedure was performed on the data for each treatment group in each section of the course. The data of common treatment groups were then summed across sections, producing a 2 X 5 frequency matrix for each category of student behavior (treatment by proportion), and a 2 X 3 frequency matrix of teacher behavior (treatment by category).

CONCLUSIONS

The results of the evaluation were most illuminating. As mentioned above, these results are summaried in Chapter III.

For the present paper our goal was the understanding of the applied evaluation process, itself. We conclude that sophisticated applied evaluations are possible in the school. The goals of the evaluator and the goals of the educator can be met simultaneously, with little serious compromise required for either. In fact, the educator and evaluator can enhance each others objectives. This conclusion is borne out by the behavior of the teachers subsequent to the evaluation. The school is actively pursuing evaluation of other aspects of the science curriculum and other departments are inquiring about how they, too, can collect such data. This facilitation of on-going evalution practices will also make it easier for other researchers to work within the school. Thus not only can the goals of the applied evaluation be met, but a style of curiosity and effective evaluation can be established in the school for the future.

Creare

References

Arons, A. B. (1984). Computer-based instructional dialogs in science courses. Science, 224, 1051-1056.

Bandura, A. (1978). The self system in reciprocal determinism. American Psychologist, 33, 344-358.

Burger, K. (1985). Computer assisted instruction: Learning style and academic achievement. Journal of Computer-Based Instruction, 12, 21-22.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design & Analysis Issues for Field Settings. Chicago: Rand McNally.

Cooley, W. W., & Lohnes, P. R. (1976). Evaluation Research in Education. New York: Irvington Publishers.

Cumming, G. (1984). Computers come to school, but psychology hasn't noticed. Australian Psychologist, 19, 275-289.

Isaac, S., & Michael, W. B. (1971). Handbook in Research and Evaluation. San Diego, CA: Edits Publishers.

Jernstedt, G. C. (1986). The relative effectiveness of computer-based and traditional instruction on science learning. Journal of Research in Science Teaching, under review.

Jernstedt, G. C., & Rothe P. H. (1986). A collaborative project to develop computer based instruction in science laboratory courses. T. H. E. Journal, in press.

Judd, C. M., & Kenny, D. A. (1981). Estimating the effects of social interventions. Cambridge, England: Cambridge University Press.

Kulik, J. A., Bangert, R. L., & Williams, G. W. (1983). Effects of computer-based teaching on secondary school students. Journal of Educational Psychology, 75, 19-26.

Kulik, C. C., Kulik, J. A., & Cohen, P. A. (1980). Instructional technology and college teaching. The Teaching of Psychology, 7, 199-205.

Papert, S. (1980). Mindstorms: Children, Computers, and Powerful Ideas. New York: Basic Books.

Rossi, P. H., Freeman, H. E., & Wright, S. R. (1979). Evaluation: A Systematic Approach. Beverly Hills, CA: Sage Publications.

Sardello, R. J. (1984). The technological threat to education. Teachers College Record, 85, 631-639.

Sawyer, T. A. (1985). Human factors considerations in computer-assisted instruction. Journal of Computer-Based Instruction, 12, 17-20.

Schunk, D. H. (1984). Self-efficacy perspective on achievement behavior. Educational Psychologist, 19, 48-58.

Stimmel, T., Connor, J. L., McCaskill, E. O., & Durrett, H. J. (1981). Teacher resistance to computer-assisted instruction. Behavior Research Methods & Instrumentation, 13, 128-130.

Wolf, F. M. (1986). Meta-analysis: Quantitative Methods for Research Synthesis. Beverly Hills, CA: Sage Publications.

Venezky, R. L. (1983). Evaluating computer-assisted instruction on its own terms. In A. C. Wilkinson (Ed.). Classroom computers and cognitive science. New York: Academic Press, pp. 31-49.

Creare

CHAPTER III: RESULTS

A Meta-Analytic Experimental Comparison of Computer-Based and Traditional Laboratory Instruction

by G. Christian Jernstedt

Though the claims and expectations for computer-based instruction (CBI) benefits in education have often been excessively great, it is clear that computer-based teaching can have positive effects on students' attitudes and learning (Kulik, Bangert, & Williams, 1983). Research that is directed at understanding the nature of the impact rather than whether it exists or not now seems to have the most appropriate emphasis. The need to understand more fully the impact of CBI is especially acute in science learning for a number of reasons. First, there is little CBI research that has been directed specifically at the science learning area. Furthermore, computers are particularly important tools in university and industry science research so there is a great deal of pressure on science teachers to integrate computers into their curriculum. The use of computers in science teaching also appears to be an especially promising way of carrying out Maehr's (1983) recommendation that science instruction help students develop independence in their learning, become confident of their abilities, and function well in environments where they can explore on their own in ways similar to the ways that professional scientists use. Finally, it is clear from the analyses of the National Assessment of Educational Progress data (Walberg, Pascarella, Haertel, Junker, & Boulanger, 1982) that classroom morale and direct instruction from the teacher are the key variables in producing positive impact in the classroom. Since CBI has been implicated in both improving morale and enhancing direct instruction, it is a potential tool for improving science education that must be more fully understood. The present study is directed at refining our understanding of computer-based instruction in the science classroom.

The impact of CBI on educational outcome can best be conceptualized in terms of a three element model of human behavior that ascribes instructional outcome to the joint effect of three domains of influence, the behavior of the student in the learning situation, the cognitive

and affective internal states that the student as a person brings to the classroom, and the environment of the classroom itself (Bandura, 1978, Jernstedt & Chow, 1980).

In work relevant to the person domain, Kulik, Kulik, and Cohen (1980) reported that attitudes and performance were not found to be directly linked in studies of CBI. In some studies attitudes improved while performance did not. In other studies performance improved but dropout rates also increased. Kulik, et al. (1983) found improvement in both performance and attitude, but the attitude results were relatively small. The impact of CBI on attitudes towards science is a matter of considerable importance, as attitudes determine future motivation towards science (Tamir, Welch, & Rakow; 1985) and there is evidence that science attitudes decrease as students in secondary schools have more exposure to science (Hofstein & Welch, 1984). Since attitudes in science courses are multidimensional, and include attitudes towards the course and its scientific discipline, the teaching methods employed in the course, and the actual class activities (Okebukola, 1985), the present study collected a variety of attitudinal data.

In the behavior domain, time on task is the principal variable that seems to predict academic achievement (Seifer & Beck, 1984). Level of task engagement has a strong impact on achievement and is, itself, influenced by teacher behavior (Tobin, 1984). Previous work, however, has failed to isolate the particular components of laboratory behavior that are most directly related to achievement gains. Okebukola (1985) reported that practical skills such as manipulation of the apparatus were important in understanding higher level, more cognitive outcomes. The present study included a set of behavioral measures of both teacher and student. These measures were specially developed to isolate the particular aspect of the laboratory activities that were most directly related to achievement. However, the direction of causal impact was not assumed to be from behavior to cognitive outcome. Rather, the behavioral and the cognitive outcomes were each considered to be separate, though probably related, components of the science learning process (Hacker, 1984).

The achievement outcomes from the science learning were also approached in the present study as multidimensional processes. In light of previous work, the performance tests were divided into three categories of performance: fact recall, use of principles, and application ability (Nicolson, Bowen, & Nicolson, 1984; Tamir, 1985).

For both practical and pedagogical reasons, computers have been employed in two different modes in science laboratory instruction. The most obvious mode is that in which students themselves use the computer, in place of traditional apparatus, to conduct experiments. In contrast to this mode of active student use of the computer is the instructor led demonstration. In this mode the instructor employs the computer, in place of traditional apparatus, to conduct a demonstration which the students observe but do not physically participate in. The literature on mode of apparatus use, whether the apparatus is a computer or not, is mixed. Some investigators find no difference between the two modes of use (Bates, 1984). However, others have reported relevant data that smaller classes are associated with higher achievement (Smith & Glass, 1980) and that the laboratory must play a central role in science teaching (Welch, Klopfer, Aikenhead, & Robinson, 1981).

Of particular importance to the mode of use of the computer is the teacher's behavior. Class size, per se, may be less important than the teacher's actual activities with students during class (Rosenshine, 1979). Practically, the demonstration mode is easiest and the least

Creare

expensive mode, since only a single computer and one trained user is required. It is likely that this simplicity and economy may encourage teachers to adopt the demonstration mode as the preferred. Therefore it is important to resolve the issue of whether one or the other mode is more effective in its impact on the educational outcome. In the current study, both modes were employed and compared. Furthermore, for the treatments where students actively used the computer, class size was systematically varied in order to separate the mode of use effects from the class size effects.

Though not directly apparent in their results, there is a suggestion in the meta-analysis conducted by Kulik, Kulik, & Cohen (1980) that the novelty of computer-based instruction, in addition to the actual instructional effect of the computer, may have influenced the quality of learning. The issue of previous experience with computers has not been directly examined for its possible impact on the use of CBI. Yet differences in degree of experience with computers are large across teachers as well as across students. The present study examined degree of experience of teacher and students for its impact on the effectiveness of both traditional and computer-based instruction.

A variable that emerged in the field notes during the pilot studies that preceded the main experiments of the present study was time of year. The behavior of the students, in terms of attention to laboratory work and motivation for studying, dropped dramatically as the academic year drew to a close. Teachers anticipated this by avoiding standard laboratory work near the end of the second semester. Since CBI influences time on task, the possibility of an interaction between CBI effectiveness and time of year must be considered. Moreover, since many of the teachers in the current study were not highly familiar at the beginning of the year with the computer methods that they were employing and had their students employing, the possibility of a novelty effect existed. Accordingly, the effects of the CBI treatment were separated in the present study into two time periods: the beginning of the year when novelty and unfamiliarity might be affecting performance, and the middle of the year, when the end of year decay had not yet reared its head and the novelty of startup had diminished.

A very unfortunate and confusing feature of much of the work on the impact of CBI has been the confounding of the media, computers, with the mode of use. Meta-analyses of different media have shown positive effects on educational achievement for most of the major media techniques that have achieved popularity (Pintrich, Cross, Kozma, & McKeachie, 1986). However, as Pintrich, et al. point out, a problem in this research has been that the method and the media of instruction have been confounded, perhaps because of reporting procedures or because it is difficult in some instances to distinguish between them. The method of using the computer was unconfounded from the media in the present study in two ways. The software developed with the computer system was specifically designed to be generic, in the sense that it did not require students or teachers to follow a fixed format or curriculum. Rather the software presented a menu of tools that could be employed in any way the user desired, so that the user, not the computer, determined the methods. Secondly, teachers and students used both traditional apparatus and computer-based apparatus in conducting their work. Each participant experienced numerous occasions of using each media. This combining of all subjects with all treatments was a major feature of the design and something that has been missing in too much of the previous work on CBI. The goal of separating the computer, as media, from the methods by which it was used appeared to be accomplished; in the laboratory, students were observed to be using, at one time or another, all four of the classifications of type of computer use that Rushby (1979) has made: instructional,

revelatory, conjectural, and emancipatory. It should be noted that the one methodology feature that was built into the computer system was to make it visually based, both because of the value of visual-based instruction (Cohen, Ebeling, & Kulik, 1981) and because the automatic graphing of data by the computer seems to be a major use and value of computers in the science classroom.

Previous work on CBI has focused far more on the computer's impact on the students' behaviors than it has attended to the teachers' behaviors. However, many classroom variables related to teacher behavior influence science achievement as much as student ability (Tamir, Welch, & Rakow, 1985). Some work indicates that CBI is effective because it allows the teacher to provide more individual attention to students (Evans, Mickelson, & Smith, 1984). One must be careful, though, not to ascribe a unidirectional impact of teacher behavior on student behavior; the impact of student and teacher behavior appears to be bidirectional (Doyle, 1979). In the present study both teacher and student behaviors were examined for their mutual impact on the educational outcome.

The goal throughout the design of the present study was to cast a wide net in an effort to detect the full richness of the processes influencing the impact of the computer on science learning. Coupled with the wide net was a conservative statistical approach to guard against overgeneralization and the quick but unreliable or invalid answer.

METHOD

Subjects

Four hundred sixty one students (209 females and 252 males) enrolled in six laboratory science courses at a regional high school participated in the study. Students in grades nine through twelve were approximately equally represented in the sample.

Treatments and Conditions

Because most students follow a regular course sequence in the science curriculum, class year was a major determinant of enrollment for each course. In addition, ability level contributed to course selection and to sectional placements within a course. The high school science department offers six courses in four basic areas: 1) a general introductory course in physical science; 2) two courses in chemistry, one general and one advanced; 3) two courses in biology, one general and one advanced; and 4) one course in physics with three levels of difficulty.

The study employed two different types of laboratory apparatus: the largely mechanical apparatus traditionally used in the classes and a computer-based educational data acquisition system. Both the computer-based and the traditional apparatus were employed in two different modes of class meetings. In 11 of the laboratories the teacher was the user of the system, giving a demonstration to the assembled class. In the other 8 laboratory sessions the students employed the computer-based system directly. For both different modes of use, the sectional divisions of each course were followed in forming group assignments for the experiments. We randomly assigned use of the computer to one to three sections and use of the traditional apparatus to one to three sections, with the total number of treatment groups dependent on the number of sections in each course. In 5 experiments, however, the length of the exercise

Creare

prohibited use of the computer for more than one lab group (4-6 students) within each section; in these instances, we randomly assigned use of the computer system to small lab groups rather than to the whole section.

The degree to which computers were used in other ways during in each course was recorded and converted to a variable indicating the combined instructor and student degree of experience with computers. Time of year was also coded, either as the beginning of the year, when teacher and students were becoming familiar with procedures and each other, or as the middle of the year when procedures had become familiar and regular but the end of year malaise and loss of attention had not yet occurred. The start of year period corresponded roughly with the first half of the first semester, while the middle of the year corresponded roughly to the second half of the first semester and first third of the second semester.

Apparatus

The computer-based educational data acquisition system (EDAS) that was used in this study is a laboratory computer system developed with funding from the U.S. Department of Education by Creare, Inc. The EDAS consists of several components, including an IBM PC dual drive computer, a Burr-Brown Data Acquisition System, and a special' constructed software package. A variety of physical sensors (pH probe, conductivity meter, pressure transducer, thermocouple, accelerometer, optical density sensor, and light source) connect to the system. The dependent variable measured by the sensor is graphed against time, with the rate and total time of sampling positioned along the x-axis of the video display. The system is modular; new sensors can be integrated by inputting the conversion ratio and other specifications of the sensor. User input is by single-character keyboard commands during data collection. The entire system is menu-driven, using directional keys that control a movable highlight.

Measurement Instruments

The organization of the measurement instruments used to obtain data about the participants in the experiments is indicated in Table III.1. Based on previous psychometric work and extensive pilot studies, a self-report inventory was constructed using 5-point scales to assess students' learning style, study habits, school activities and attitudes, personality (including motivation, self-esteem, and social behavior). The inventory also collected other demographic information. In addition, the inventory asked students to evaluate their course on a number of dimensions, including instructor characteristics, particular features of the course (e.g., lectures, labs, audio-visual instruction, computers used in lab ·, and overall quality. The original version of this survey was composed of 160 questions; this form was administered during the pilot-testing phase (the spring semester preceding the 85-86 academic year) and factored to the current form, which consists of 110 questions. Three versions of the inventory were established, all of which were equivalent in form, length, and questions asked, but which differed in terms of timing. One version asked students to answer according to what they expected their science course to be like for the current year; one asked students to answer according to what they had experienced in the science course during the first semester; and one asked students to answer according to what they had experienced in the science course during the second semester.

Creare

Table III.1

Measurement Instruments

Domain	Source of information	Interval of measurement	Variable being measured
Person	Student	Semester	Learning style Study habits School activities School attitudes Personality Demographics
Person	Student	Laboratory	Attitudes towards class Enjoyment Difficulty Interest Attitudes towards computers Enjoyment Comfort with Motivation To learn To do well in course Attitudes towards lab Knowledge gained Confidence in abilities Effectiveness of methods Know what are doing Independence Rushed
Person	School	Year	Courses taken Grades in courses

Creare

Table III.1 (continued)

Domain	Source of information	Interval of measurement	Variable being measured
Behavior	Observer	Laboratory	Off-task Group involvement Individual attention from teacher Attention to teacher by group Attention to the apparatus Writing / note taking
Behavior	Teacher	Laboratory	Lab report Fact recall Graphing ability Use of principles Application ability Total score Examination Fact recall Use of principles Application ability Total lab question score Total exam score

Table III.1 (continued)

Domain	Source of information	Interval of measurement	Variable being measured
Environment	School	Year	Demographics
Environment	Teacher	Year	Course characteristics
Environment	Student	Year	Student expectations for the course
Environment	Student	Semester	Course characteristics Instructor characteristics
Environment	Experimenter	Laboratory	Science discipline Degree of computer experience Time of year Mode of use student active use instructor led demonstration

Creare

A daily report questionnaire was also constructed to assess pre- and post-class attitudes in the laboratory sessions. As with the inventory, these daily reports employed 5-point scales, where one extreme represented full agreement and the other extreme represented no agreement. On these daily reports, students were asked to rate their enjoyment of the class, their interest in the subject matter, the difficulty of the class, their enjoyment of using computers and their degree of comfort with using computers. In addition, two unique questions were asked on the pre-class form to assess before-treatment motivation to learn the material and to do well in class. Six such unique questions were asked on the post-class survey: the amount of knowledge students felt they had learned in the class; the students' degree of confidence in their abilities during class; the effectiveness of the methods used to learn in class; the degree to which students' felt they knew what they were doing in class; the degree to which students followed directions or worked independently; and the extent to which students felt rushed during the class period. Teachers filled out a similar form after the class meeting reporting on what they thought class was like for the students.

An observational checklist for the assessment of teacher and student behaviors was developed with extensive pilot work and used in the individual laboratory sessions. Three categories of teacher behaviors were included in the checklist: off-task to all students (e.g., out of room); providing individual attention to students; and providing attention to groups of students (e.g., lecturing). Observations of teacher behaviors were recorded at one-minute intervals. The three categories of teacher behaviors were considered mutually exclusive; only the single category most displayed during an observational period was recorded. Six types of student behavior were included in the checklist: off-task; on-task group involvement in laboratory activities; receiving individual attention from the instructor; attending as a group to the instructor; attending to the apparatus; and note-taking or other forms of writing. Estimates of the proportion of students engaged in each category of behavior were recorded after each successive minute of class time. Types of student behavior were not considered unique; the same student who engaged in two types of behavior during a single interval was counted in the total proportion of each category. In other words, the proportions of students engaged in all categories of behaviors did not sum to one. There was considerable overlap between categories. For instance, students who were receiving individual attention from the instructor were sometimes writing at the same time. However, when the instructor was on-task to the entire class, as during a lecture, no students were construed to be individually engaged with their teacher. Five proportions of class participation were used by the observers: 0, representing no students; 0.25, representing from 1 student to 1/4 of the sample; 0.5, representing more than 1/4 of the sample but not more than 1/2; 0.75, representing more than 1/2 of the sample but not more than 3/4; and 1, representing more than 3/4 of the sample. One highly trained experimenter recorded all behavioral observations.

Assessment of academic performance was made by obtaining teacher-graded reports, quizzes, and examinations. No special performance measures were constructed; only the assignments included in the normal grading process were analyzed. Data prepared for analysis included scores on each question of a particular exercise that related to the students' work during the experimental lab sessions, categorized into fact recall, use of principles, or applications of material to new situations. Subscales formed by summing the scores on all laboratory related questions and by summing all scores on the total exam were also created for each exercise.

Creare

Procedure

Experiments were conducted over the full academic year 1985-86. The self-report inventories were administered in each course during the first two weeks of the first semester, in the first two weeks of the second semester, and in the final two weeks of the academic year. These inventories were administered in individual class meetings rather than in large group meetings of all sections of the course. For all administrations the experimenter read standard instructions, part of which cited concrete examples of how student responses were being taken seriously and were making a difference in the way the courses were taught and organized.

Experimental comparisons of the two laboratory instructional techniques were conducted throughout the academic year. The regular topical sequences of each course and experiments were followed and experimental comparisons were conducted only with regularly scheduled laboratory exercises. To avoid detrimental novelty effects due to teacher unfamiliarity with the computer system, we worked in conjunction with the teachers in planning and pilot-testing those laboratories in which use of the computer was possible. The computer was only employed for laboratories in which three criteria were met: (a) the system could sense the process in question, (b) the variables involved could be graphed against time, and (c) the system could be used in a way analogous to the way the traditional apparatus was used in demonstrating the process or principle in question.

The computer-based system was compared to the traditional apparatus in 19 laboratories during the course of the year. A standard procedure was followed in each experiment. After attendance was taken by the teacher, the experimenter addressed the class, indicating that as part of the on-going evaluation of their science course, the students were being asked to fill out two questionnaires in that day's class, one at the beginning and one at the end. Since the experimenter's presence may have been associated with the computer, students may have expected to use the system in the test particular sessions. In order to prevent expectancy effects from distorting the pre-class daily report, the computer was present in the room during the introductions for all sections and in plain view of all students. To further minimize expectancy effects among the students, during the year the experimenter also administered the questionnaires in class meetings in which there were no experimental comparisons being conducted. Once all students were finished with the pre-class report, the experimenter assigned students to their lab group, and then proceeded to record observations of teacher and student behaviors. In those labs in which both instructional media were present, the experimenter alternated recording observations for minute-long periods between the computer group and the control group. When the first lab group finished the procedures, observations were terminated and the post-class questionnaires were handed out to each group as they finished.

Design and Data Analyses

The evaluation was structured as a series of 19 independent experiments, each with two treatments: the experimental (EDAS computer-based) and the control (traditional apparatus). Variables were clustered in three predictor domains and one criterion domain as indicated in Figure III.1.

Each experiment was first analyzed separately with a complete multivariate analyses. All Person, Environment, and Outcome variables were tested with a multivariate analysis of

variance. Then each member of the domain was examined with univariate analyses of covariance, using a general linear model. Factor analyses and regression analyses during the pilot phase of the evaluation indicated that a student's grade in the science course being studied was the most appropriate indicator of ability, thus it was used as the covariate. Early in the year recent grades in the science course served as covariate. Once a quarter course grade was available it served as covariate. When semester grades became available, they were used as the covariate. When the analysis of covariance showed no significance for the covariate, univariate analyses of variance were employed.

Creare

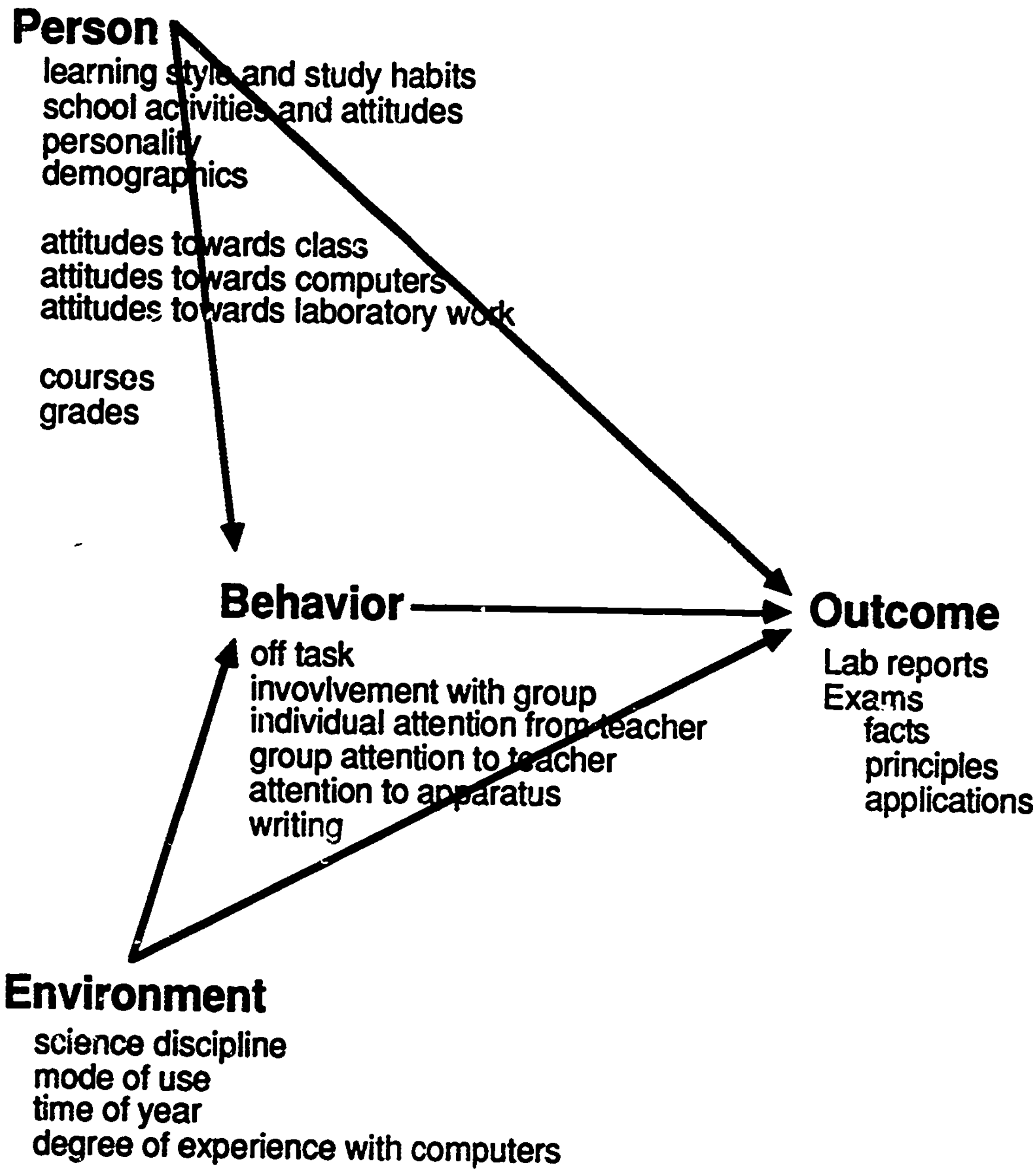


Figure III.1. The domains of the variables and the sets of variables that were examined during the experiments for their impact on the educational outcome.

Creare

The behavioral data were obtained as frequency of occurrence for each of the six categories of student behavior and three categories of instructor behavior. These data were analyzed with chi-square. When there were cells that had too few observations to meet stringent chi-square criteria, frequencies were collapsed across cells.

After all 19 experiments had been analyzed independently with the multivariate procedures, meta-analysis techniques were employed to examine the data from all experiments simultaneously. For all variables but the behavioral, effect sizes were computed directly from the raw data in each experiment. For the behavioral variables, the effect sizes were computed from the chi-square results. To detect possible interactions between treatments under different conditions, the meta-analyses employed coding, or indicator, variables. The major environmental variables that served as indicators were: (a) mode of use of the apparatus, whether actively by the students or in an instructor led demonstrations, (b) degree of experience with computers, which varied from courses where the EDAS was the only computer in use to courses where a variety of computers were used for a variety of different purposes by both students and instructor, and (c) time of year, varying between the start of the year when novelty effects could still be occurring and the middle of the year when attention to school work was maximal for both instructor and student and all procedures and apparatus was familiar. Pearson product moment correlation coefficients were obtained between each indicator and the effect sizes for each variable. Where interactions between indicators were possible, the correlations were obtained by examining each level of each indicator, in turn, and computing a correlation matrix between the other indicators and the variable effect sizes.

RESULTS

The computer-based treatment differed from the traditional in its effect on variables in all of the domains. The meta-analyses indicated that these effects showed regularity across the different experiments. Each of the indicator variables was associated with significant differences in the effect of the two treatments.

Both the huge amount of data to be summarized and the coherence of the picture that emerges from the analyses make a tabular presentation of the results most effective. Accordingly, Tables III.2 through III.5 summarize the findings. With effect size computations in meta-analyses, reporting the actual means averaged across experiments can be misleading and invalid, as the difference between averages of means is not directly proportional to the effect size and may not even have the same sign. Therefore, only representative means and variance analyses will be reported in Table III.5. Tables III.2 through III.4 present the meta-analysis findings, containing both the correlation coefficients (r) between the indicators and the effect sizes and the effect sizes (d), themselves, for all significant meta-analytic results.

In Table III.2 the meta-analysis results are presented for each of the indicators considered separately. For the indicator, time of year, the computer group changed from reporting more comfort with computers to reporting less comfort than the traditional group. With enjoyment of computers the computer group reported less enjoyment than the traditional at mid year. Off task behavior for the computer group changed from more than that for the traditional at the start fo the year to less than the traditional by mid year. Attention to the apparatus showed an opposite effect. Amount of writing became lower for the computer group than the traditional by midyear. Lab report scores, which were higher for the computer groups at the start of the year, showed no difference across groups at midyear.

Creare

For the indicator, mode of use, the computer students in the active use groups reported less interest in the course, a less difficult class, and a less rushed class than did the traditional. Writing behavior was less for the computer students than for the traditional in the active use classes, but exam application scores were higher.

For the degree of experience indicator, when experience was high the computer users reported less confidence in their abilities than did the traditional. The behavior of the computer groups showed more off task activities than that for the traditional. Exam performance on both principle and application questions, as well as on the sum of all laboratory-relevant questions was higher for the computer groups than for the traditional.

Creare

Table III.2

Global Meta-Analysis Results For Each Indicator

Time of Year (1 = start, 2 = middle)

D	Variable	r_1	p	r_2	p	d_1	d_2	N
P	Comfort with computers	.48	<.05	-.76	<.001	.28	-.34	19
P	Enjoyment of computers			-.69	<.001	.30	-.42	19
B	Off task behavior	.56	.01	-.57	.01	.33	-.55	19
B	Attention to apparatus	-.53	<.05	.56	.01	-.28	.45	19
B	Writing			-.56	.01	-.14	-.92	19
O	Lab report: total score	.85	<.10			1.04	.00	19

Mode of Use (1 = student, 2 = demonstration)

D	Variable	r_{12}	p	d_1	d_2	N
P	Interest in course	-.44	<.06	-.21	.07	19
P	Difficulty of class	63	<.01	-.47	.09	19
P	Rushed class	-.61	<.01	-.44	.37	19
B	Writing	-.65	<.01	-.88	-.03	19
O	Exam: applications	.67	.001	.39	-.22	19

Creare

Table III.2 (continued)

Degree of Experience (1 = inexperienced, 2 = experienced)

D	Variable	r_{12}	p	d_1	d_2	N
P	Confidence in ability	-.47	<.05	.17	-.25	19
B	Off task behavior	.46	<.05	-.27	.25	19
O	Exam: lab questions	.60	<.01	-.04	.28	19
O	Exam principle questions	.50	<.10	-.02	.32	19
O	Exam: application questions	.68	.001	-.22	.39	19

Note. D is the domain from which the variables are sampled: where P is the person, B is the behavior, E is the environment, and O is the outcome domain. The number of experiments that d_1 and d_2 are each based upon is indicated by N. When a correlation is not significant, it is omitted from the table and the corresponding effect size is obtained from the correlation for the opposite level of the indicator.

Creare

The effects of the mode of use also differed across levels of the other indicator variables, as presented in Table III.3. Consider the mode of active student use first. When the laboratory apparatus was employed by small lab groups rather than by the whole lab section, computer groups reported more motivation for class and experienced less time receiving individual attention from the teacher than did traditional sections. The time of year indicator revealed similar effects as it did across all modes of instruction with the exception that group attention from the teacher was less by midyear for the computer groups than for the traditional. Comparing courses with high computer experience to those with little computer experience revealed student reports for the experienced computer users of lower interest in class, less confidence in their ability, and less sense of knowing what they were doing, and observer reports of more group attention from the teacher than was found with the traditional apparatus users

Table III.3

Meta-Analysis Results For Indicator: Mode of Computer Use

Student active use

Indicator	D	Variable	r	p	d_1	d_2	N
Size of lab group	P	Motivation for class	.75	<.05	-.27	.23	8
	P	Individual attention	-.71	<.05	.47	-.17	8
Time of year		see Table 2 also					
	B	Attention to teacher	-.68	<.10	.29	-.23	8
Degree of experience	P	Interest in class	-.71	>.05	.19	-.53	8
	P	Confidence in ability	-.69	<.10	.41	-.34	8
	P	Know what doing	-.75	<.05	.37	-.34	8
	B	Attention to teacher	.68	<.10	-.23	.29	8

Instructor led demonstration

Indicator	D	Variable	r	p	d_1	d_2	N
Time of year		see Table 2 also					
	P	Confidence in ability	-.74	<.01	.40	-.10	11
	P	Know what doing	-.62	<.05	.32	-.18	11
	B	Attention to teacher	.54	<.10	-.41	.08	11
	B	Writing	not significant				
	O	Exam: applications	.57	<.10	-.31	-.05	11
	O	Exam: total	.53	<.10	-.04	.40	11

Creare

Table III.3 (continued)

Instructor led demonstration

Indicator	D	Variable	r	p	d_1	d_2	N
Degree of experience	P	Motivation for course	.53	<.10	-.02	.54	11
	P	Interest in class	.55	<.10	-.15	.33	11
	P	Learned in class	.61	<.10	-.01	.53	11
	P	Difficulty of class	.60	<.10	-.20	.57	11
	B	Off task behavior	.77	<.01	-.20	.71	11
	B	Group involvement	.64	<.05	-.05	.23	11
	B	Attention to teacher	-.79	<.01	.10	-.73	11
	B	Attention to apparatus	-.78	<.01	.10	-.73	11
	O	Exam: lab questions	.67	<.05	-.07	.34	11
	O	Exam: principles	.63	<.05	-.01	.34	11

Note. The indicator is the coding characteristic of the laboratory class meeting that is employed in the analysis to create two categories for comparison. These indicator paired-comparisons are (a) size of lab group: whole class vs small groups, (b) time of year: start vs middle, and (c) degree of experience with computers: little vs much. D is the domain from which the variables are sampled: P = person, B = behavior, E = environment, and O = outcome. The effect size for the first named member of the paired comparisons is d_1 and for the second named member of the pair, d_2. The number of experiments that d_1 is based upon is indicated by N. The sum of the two N's for d_1 and d_2 is 19. For time of year, the statistics are, whenever available, averages of those for the separate analyses of start vs rest of year and middle versus rest of year.

Creare

Consider next the effects obtained when teachers led demonstrations with the laboratory apparatus. The effects of time of year were similar to those across all modes of use with a few additions. By midyear the increased confidence and knowing what they were doing, relative to the traditional groups, was lost for the students in the computer groups. The computer groups received more group attention from the teacher than the traditional by midyear. Finally, the application questions on the exams became equal across conditions by midyear, while the total exam score for the computer groups rose above that for the traditional at midyear. In courses with high computer experience demonstrations resulted in the following changes for the computer groups when compared to the traditional. The computer groups showed more motivation, more interest in class, and reported that they learned more, though the class was more difficult than for the traditional groups. Off task and group involvement behaviors increased for the computer groups when compared to the traditional, while attention to the teacher and the apparatus decreased relative to the traditional. Exam scores on both the principle questions and the sum of all questions directly relevant to the laboratory exercises rose for the computer groups compared to those for the traditional.

The impact of the degree of experience with computers for teacher and students varied across levels of the other indicator variables, as well, and is presented in Table III.4. Consider the courses with experienced users first. The time of year effect was similar to that previously reported with a few additions. By mid year the difficulty of the course increased and the independence of the laboratory activities decreased for the computer groups when compared to the traditional. On the exams both the principle and the application question scores were higher for the computer than for the traditional at midyear.

Creare

Table III.4

Meta-Analysis Results For Indicator: Degree of Computer Experience

Experienced with computers

Indicator	D	Variable	r	p	d_S	d_M	N
Time of year		see Table 2 also					
	P	Difficulty of course	.75	<.10	.00	.84	6
	P	Independence in lab	-.91	<.01	-.05	-2.18	6
	B	Off task behavior	not significant				
	B	Attention to apparatus	not significant				
	O	Exam: principles	.82	<.05	.19	1.00	6
	O	Exam: applications	.98	<.001	.03	1.00	6

Creare

Table III.4 (continued)

Inexperienced with computers

Indicator	D	Variable	r	p	d_S	d_M	N
Time of year		see Table 2 also					
	P	Enjoyment of course	not significant				
	P	Learr .. in class	.53	<.10	-.55	-.04	13
	B	Off task behavior	not significant				
	O	Lab report: total	not significant				
	O	Exam: total	.61	<.05	-.04	.38	13
	O	Exam: applications	.59	<.05	-.31	-.04	13

Note. The indicator is the coding characteristic of the laboratory class meeting that is employed in the analysis to create two categories for comparison. For time of year the two categories are start vs middle of year. D is the domain from which the variables are sampled: P = person, B = behavior, E = environment, and O = outcome. The effect size for the start of the year is d_S and for the middle, d_M. The number of observations that d_S is based upon is indicated by N. The sum of the two N's for d_S and d_M is 19. The statistics are, whenever available, averages of those for the separate analyses of start vs rest of year and middle versus rest of year.

Creare

Earlier results can be restated for experienced users with demonstrations as follows. The computer users, when compared with traditional apparatus users, reported for demonstrations more interest in the course and in class meetings, more sense of learning, but greater difficulty, and showed less attention to the teacher than they reported and showed when they were in active use of the apparatus.

Next consider the users with less experience. Time of year effects were, again, similar to the overall results of time of year with a few exceptions. The differences in enjoyment of computers, off task behavior, and lab report scores were not present. Changes not revealed in the overall analysis included for the computer groups less sense of learning in class at the start of the year than was felt by the traditional groups. In addition, the computer groups performed worse on the total exam at the beginning of the year but better by midyear than did their traditional counterparts. A low exam application score for the computer groups at the start of the year disappeared by midyear when both computer and traditional groups were equal on this measure.

Earlier results can be restated for inexperienced users with demonstrations as follows. With demonstrations rather than active student use the computer groups reported less interest, confidence in their ability, feelings that the methods used were not effective, and sense of knowing what they were doing, while reporting more difficulty and sense of being rushed than the traditional group members reported. Group involvement and attention to the apparatus were also less, while attention to the teacher and writing were greater for these computer users than for the traditional apparatus users.

Creare

Table III.5

Representative Means and Analysis Results

Person domain

Variable	Time Period	Means Computer	Means Traditional	F	df	p
Difficulty of class	pre + post	2.60	2.10	3.98	1,28	.05
Enjoyment of computers	pre x post	2.34 3.16	3.29 3.21	12.13	1,42	.001
Comfort with computers	pre + post	4.08	4.42	6.41	1,57	.01
Confidence in ability	post	3.65	5.35	5.74	1,22	.02
Know what doing	post	3.36	4.10	6.32	1,32	.02
Motivated	pre	3.33	2.60	5.96	1,28	.02
Interest in class	pre + post	2.43 3.00	2.59 2.68	6.16	1,34	.02
Amount learned	post	3.53	3.15	4.65	1,62	.02
Rushed	post	2.15	1.59	4.58	1,58	.03
Effectiveness of methods	post	2.79	3.50	4.03	1,32	.05
Independence	post	3.53	4.14	4.05	1,22	.05

Creare

Table III.5 (continued)

Behavior domain

Variable	Sign of d	Means Computer	Traditional	X^2	df	p
C ff task	+	.292	.116	59.74	1	<.0001
	-	.119	.196	13.84	1	.0003
Group involvement	+	.72	.65	8.56	2	.01
	-	.546	.597	4.14	1	.04
Individual attention	-	.120	.161	8.39	2	.02
Attention to teacher	+	.222	.139	26.44	2	<.0001
	-	.155	.216	9.67	2	.008
Attention to apparatus	+	.960	.897	13.01	2	.00?
	-	.642	.748	14.04	1	.0003
Writin	+	.126	.064	12.14	2	.0?3
	-	.253	.352	8.57	1	.004

Creare

Table III.5 (continued)

Outcome domain

Variable	Time Period	Means Computer	Means Traditional	F	df	p
Lab report: total	post	10.00	7.17	12.98	1,30	.001
Exam: total	post	10.00	8.77	4.06	1,57	.05
Exam: lab questions	post	10.00	7.59	4.85	1,57	.03
Exam: principles	post					
Exam: applications	post	10.00	3.11	6.55	1,57	.01

Note. For each row of the table with the person and outcome variables an analysis of covariance or variance is reported that represents the difference score results of the earlier meta analyses. The time period indicates whether the variable was measured just before (pre) or at the end (post) of class. When only one period is presented, only that period was significant. When both periods are summed, only the main effect of treatment was significant. When the two periods are multiplied, the interaction over time was significant and is presented in the table next to the post scores. All attitude scales range from a low score of 1 to a high of 5. Outcome measures have been standardized so that the experimental group score is set at 10. For the behavior variables a chi-square analysis is reported that represents the difference score results of the earlier meta-analyses. Behavioral scales are proportions of time in which the class was observed to be engaged in that behavior. The sign of d column reports whether the effect size indicated difference in favor of the computer groups (+) or the traditional groups (-). Where they existed, results for both directions of effect sizes are reported.

Creare

DISCUSSION

A complex though comprehensible picture emerged of the multidimensional and multidomain nature of educational outcome from CBI in science teaching. The conflicting changes in behavior and attitude found by Kulik, Kulik, & Cohen (1980) and the small attitude effects reported by Kulik, et al. (1983) are understandable in light of the present results. Most interesting were the differences between attitudes and performance. In some cases, such as for courses with a high degree of computer experience where the demonstration mode was used, improved attitudes were directly proportional to increased achievement. Yet when demonstrations were compared with active student use of the computer, the demonstration mode resulted in improved attitudes but decreased achievement. Our knowledge has advanced to the stage where research on CBI must search for the confounding variables and seek to understand the complexity of the underlying process rather than simply demonstrate a difference in scores.

The particular variables isolated in the design of the study were found to play important roles in determining the impact of the CBI. Time of year, degree of computer experience, and mode of use each produce main effects on outcome and interact with other variables in significant ways. No single condition is optimal. Instructors must choose tools and approaches on the basis of their goals; they cannot, unfortunately, simple choose the universally best treatment.

One rather general effect was the change in attitudes towards computers, the course, and themselves that students reported with increased exposure to computers. When computer exposure increased, whether through passage of time during the academic year, by being in a class with extensive computer use, or by using the computer actively rather than observing a demonstration, students reported less liking of and comfort with computers. Often under these same conditions students also reported less confidence in their own abilities and a diminished sense of knowing what they were doing and thinking that the methods being used in class were effective. It is important to remember that these decreases were not from positive to negative feelings, but were from very positive to somewhat less positive (from 4.4 to 4.0 on a 5 point scale, for example). Field observations suggested two related reasons for this change. With more computer experience students became more realistic in their expectations of what could be done with the computer. Initially many thought that the introduction of the computer would dramatically simplify their task in lab. As the year progressed they discovered that, while the clerical work was decreased, there still was plenty to do in lab. Furthermore, both students and teachers found that the use of the computer did not make learning simpler in a cognitive sense. In fact, both found that the computer made learning seem more complex in many ways. By eliminating much of the writing and other clerical tasks, the computer freed students and teacher to think more about the content of what they were doing. The ease and speed of collecting data meant that there was more time to explore more questions in lab. Both the teachers and students found this increase in cognitive responsibility to be harder and more challenging. In contrast to the simple, almost rote activities of earlier lab experiences students found science to be hard work, especially when the new tool clearly removed the rote activities. This suggests that the computer accelerated the contrast between simple and more professional science that has been reported by Hofstein and Welch (1984) as occurring between junior high school and high school.

Creare

Concurrent with this decrease in attitudes was an increase in performance. Later in the academic year CBI students engaged in less time off task during the laboratory and shifted their time from writing down data to attending to the apparatus. This improvement in behavior was not directly linked to an improvement in their laboratory reports; the computer groups initially wrote better reports than the others, but this difference disappeared by the middle of the year. For classes with high computer experience there also emerged an improvement in examination performance by mid year, but this improvement did not occur as clearly for the classes with less teacher and student experience. Future work must carefully distinguish between experienced and inexperienced classes when examining CBI.

The mode in which the computer was used significantly influenced the impact of the computer on students. One of the contrasts between the modes of use was that between experienced and inexperienced classes. The results indicate that in the inexperienced classes the demonstration mode was less successful than the active student use mode. Field observations suggested that this effect was due to the experience of the teacher more than that of the students. Teacher resistance to CBI has been documented by previous work (Stimmel, Connor, McCaskill, & Durrett, 1981). The behavior of the teacher can also influence the level of cognitive activity that occurs in the science classroom (Hacker, 1984). When using the computer in a demonstration mode the teacher's attitudes and experience are most exposed to the students. Inexperienced teachers were less confident and more confused in demonstrations than were the experienced teachers. The attitudes and behaviors of the students seemed to reflect these differences, though the students' performance on examinations was not affected.

At the beginning of the year students expected a great deal from their computer demonstrations, but this expectation was lost by midyear. The behavioral and outcome data do not correspond to this change in attitudes, especially for classes with computer experience. In fact, the behavioral data appear at first to be confusing for these classes. While motivation and interest was higher, there was more off task behavior and less attention to the teacher and the apparatus. This anomaly can be understood by examining the group involvement behavior, which increased with experience. Classroom observation revealed this increase as resulting from greater student discussion about the process being demonstrated. While the discussion was associated with less traditionally expected measures on task behavior, it, nevertheless, was related to improved ability to remember the relevant principles and apply them to new situations.

High school science laboratory exercises are traditionally used both to familiarize students with the methods and apparatus of science and to provide immediate examples of physical principles and relationships. In the first of these purposes the science lab exercise may very well succeed to the extent that precision is maintained in the procedures. In its second purpose, the science lab may not be serving directly and immediately to add much to the learning of facts and concepts. Data taken during class remains in tabular form until students are required to produce a graph as a homework exercise. While those with steady study habits may be prompt in completing such exercises, for most the principle or relationship which the data conveys remains hidden until the lab is discussed in a later class period. It is only when the data take on a specifically meaningful form that they begin to be organized in easily recallable fashion. It therefore stands to reason that the more the data assume the necessary form to convey the lab principle at the time of collection, the longer the students have to practice with this form of the data.

Creare

One might expect, then, that the active student labs would be more effective than the demonstrations. The actual picture is not quite that simple. Student labs were seen as less interesting than demonstrations, but were also reported to easier and less rushed. Writing behavior was also less for student labs. Apparently the use of the computer did eliminate some of the clerical work load. Key to understanding the student labs are the examination data. Application ability improved for students when they actively used the computer in class. There is some indication that the learning strategies differed in the active student labs (see Weinstein & Mayer, 1986, on strategies), and this resulted in an increased ability to use more broadly what was learned. It is important to note that active student use of the computer was not generally effective for all types of examination performance; its effect was focused on the ability to apply.

The results of the 19 experiments do have one common pattern to them. Technology can provide tools for learning, but it is what people do with the tools that determines their effect. As argued by White and Tisher (1986), curriculum is a matter of people. The present data have confirmed that what computers do is provide a potential technique for helping students develop abstract reasoning skills and an enhanced understanding of science (Arons, 1984). The results further suggest how this potential is converted into reality.

Creare

References

Arons, A. B. (1984). Computer-based instructional dialogs in science courses. Science, 224, 1051-1056.

Bandura, A. (1978). The self system in reciprocal determinism. American Psychologist, 33, 344-358.

Bates G. C. (1978). The role of the laboratory in secondary school science programs. In M. B. Rowe (Ed.), What research says to the science teacher. (Vol.1, pp. 55-82). Washington, DC: National Science Teachers Association.

Cohen, P. A., Ebeling, B. J., & Kuhik, J. A. (1981). A meta-analysis of outcome studies of visual based instruction. ECTC, 29, 26-36.

Doyle, W. (1979). Classroom tasks and student abilities. In P. L. Peterson & H. L. Walberg (Eds.), Research on teaching: Concepts, findings, and implications. (pp. 183-209). Berkeley, CA: McCutchan Publishing Coporation.

Evans, P., Mickelson, N., & Smith, N. C. (1984). Computer-assisted instruction in college composition. Alberta Journal of Educational Research, 4, 311-319.

Hacker, R. G. (1984). A hierarchy of intellectual development in science. British Journal of Educational Psychology, 3, 5-17.

Hofstein, A., & Welch, W. W. (1984). The stability of attitudes toward science between junior and senior high school. Research in Science and Technological Education, 2, 131-138.

Jernstedt, G. C., & Chow, W. K. (1980). Lectures and Textual Materials as Sources of Information for Learning. Psychological Reports, 46, 1327-1339.

Kulik, J. A., Bangert, R. L., & Williams, G. W. (1983). Effects of computer-based teaching on secondary school students. Journal of Educational Psychology, 76, 19-26.

Kulik, C. C., Kulik, J. A., & Cohen, P. A. (1980). Instructional technology and college teaching. The Teaching of Psychology, 7, 199-205.

Maehr, M. L. (1983). On Doing Well in Science: Why Johnny No Longer Excels; Why Sarah Never Did. In S. G. Paris, G. M. Olson & H. W. Stevenson (Eds.), Learning and Motivation in the Classroom. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Nicolson, R. I., Bowen, P., & Nicolson, M. K. (1984). Classroom evaluation of 'SUMS' CAL arithmetic program. Human Learning: Journal of Practical Research, 3, 129-136.

Okebukola, P. A. (1985). Effects of student-student interactions on affective outcomes of science instruction. Research in Science and Technological Education, 3, 5-17.

Creare

Ckebukola, P. A. (1985). Science laboratory behavior stategies of students relative to performance in and attitudes toward laboratory work. Journal of Research in Science Teaching, 22, 221-232.

Pintrich, P. R., Cross, D. R., Kozma, R. B., & McKeachie, W. J. (1986). Instructional Psychology. In M. R. Rozenweig & L. W. Porter (Eds.), Annual Review of Psychology, 37, 611-651.

Rosenshine, B. V. (1979). Content, time, and direct instruction. In P. L. Peterson & H. L. Walberg (Eds.), Research on teaching: Concepts, findings, and implications. (pp. 183-209). Berkeley, CA: McCutchan Publishing Corporation.

Rushby, N. (1979). An Introduction to Educational Computing. London: Croom Helm.

Seifer, E. H. & Beck, J. J., Jr. (1984). Relationships between task time and learning gains in secondary schools. Journal of Educational Research, 78, 5-10.

Smith, M. L. & Glass, G. V. (1980). Meta-analysis of research on class size and its relationship to attitudes and instruction. American Educational Research Journal, 17, 419-433.

Stimmel, T., Connor, J. L., McCaskill, E. O., & Durrett, J. (1981). Teacher resistance to computer-assisted instruction. Behavior Research Methods and Instrumentation, 13, 128-130.

Tamir, P. (1985). Meta-analysis of cognitive preferences and learning. Journal of Research in Science Teaching, 22, 1-17.

Tamir, P., Welch, W. W., & Rakow, S. J. (1985). The influence of science class attitudes and teacher image on student outcomes. Journal of Research and Development in Education, 3, 5-17.

Tobin, K. (1984). Student task involvement in activity oriented science. Journal of Research in Science Teaching, 21, 469-482.

Walberg, H., Pascarella, E., Haertel, G., Junker, L., & Boulanger, D. (1982). Probing a model of educational productivity in high school science with national assessment samples. Journal of Educational Psychology, 74, 295-307.

Weinstein, C. E. & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittlock (Ed.), Handbook of research on teaching: A Project of the American Educational Research Association (3rd ed.). (pp. 315-325). New York: MacMillan Publishing Company.

Welch, W. W., Klopfer, L. E., Aikenhead, G. S., & Robinson, T. (1981). The role of inquiry in science education: Analysis and recommendations. Science Education, 65, 25-31.

Whte, R. T. & Tisher, R. P. (1986). Research on Natural Sciences. In M. C. Wittlock

(Ed.), Handbook of research on teaching: A Project of the American Educational Research Association (3rd ed.).(pp. 874-905). New York: MacMillan Publishing Company.

Creare Incorporated is an advanced engineering consulting organization and a computer products company that provides services and products to an international clientele. Founded in 1961, Creare's business is concentrated in four areas: engineering research and development consulting, communications and graphics software, data acquisition systems and computational fluid dynamics.

Creare Inc. • Etna Road • P.O. Box 71 • Hanover, NH 03755 • 603-643-3800